

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 7, 2003, 19:15:54 ; Search time 65 seconds
(without alignments)
1467.330 Million cell updates/sec

Title: US-09-820-425B-12

Perfect score: 311

Sequence: 1 actccagctctgtgtgcaag.....gccatttcaggtaaagctt 311

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 441362 seqs, 153338381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents_NA.*

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PCFUS_COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	5.8	927	4	US-09-134-001C-1597
2	17	5.5	2163	4	US-09-480-921B-7
3	17	5.5	5077	4	US-09-480-921B-24
4	16	5.1	423	4	US-08-651-155B-56
5	16	5.1	831	4	US-09-118-55A-66
6	16	5.1	831	4	US-09-118-627-66
7	16	5.1	831	4	US-09-602-877A-66
8	16	5.1	844	4	US-09-889-595-2
9	16	5.1	1011	1	US-08-332-312-3
10	16	5.1	1282	3	US-08-961-083-171
11	16	5.1	1669	4	US-09-026-039-1
12	16	5.1	2730	2	US-08-811-897A-39
13	16	5.1	2730	2	US-08-855-213-39
14	16	5.1	2730	4	US-09-201-474-39
15	16	5.1	2814	2	US-08-811-897A-38
16	16	5.1	2814	2	US-08-855-213-38
17	16	5.1	2814	4	US-09-201-474-38
18	16	5.1	3144	4	US-08-961-527-224
19	16	5.1	5058	4	US-09-889-595-1
20	16	5.1	5852	4	US-09-853-768-10
21	16	5.1	7037	4	US-09-853-768-3
22	16	5.1	10660	2	US-08-267-803B-8
23	16	5.1	10660	4	US-09-041-886-16
24	16	5.1	30549	4	US-09-134-001C-322
25	15	4.8	20	4	US-09-323-743-53
26	15	4.8	20	4	US-09-323-743-54
27	15	4.8	20	4	US-09-323-743-55

c 28	15	4.8	156	1	US-08-392-678-26	Sequence 26, Appl
c 29	15	4.8	156	1	US-08-457-304A-26	Sequence 26, Appl
c 30	15	4.8	156	1	US-08-456-701A-26	Sequence 26, Appl
c 31	15	4.8	156	4	US-08-684-932A-26	Sequence 412, App
c 32	15	4.8	609	4	US-09-328-111-412	Sequence 168, App
c 33	15	4.8	683	4	US-09-328-111-168	Sequence 1, Appl
c 34	15	4.8	799	3	US-09-095-485-1	Sequence 24, Appl
c 35	15	4.8	800	1	US-08-368-803-24	Sequence 25, Appl
c 36	15	4.8	800	2	US-08-578-096A-25	Sequence 25, Appl
c 37	15	4.8	800	3	US-09-240-426-25	Sequence 5, Appl
c 38	15	4.8	926	1	US-08-081-448-5	Sequence 6, Appl
c 39	15	4.8	926	2	US-08-470-670A-6	Sequence 1, Appl
c 40	15	4.8	926	3	US-08-481-739-1	Sequence 1, Appl
c 41	15	4.8	926	4	US-09-167-921-1	Sequence 39, Appl
c 42	15	4.8	926	4	US-09-277-020-39	Sequence 1, Appl
c 43	15	4.8	926	4	US-09-323-743-1	Sequence 6, Appl
c 44	15	4.8	926	4	US-08-461-511A-6	Sequence 5, Appl
c 45	15	4.8	926	4	US-09-271-014A-5	

ALIGNMENTS

RESULT 1

US-09-134-001C-1597
; Sequence 1597, Application US/09134001C
; Patent No. 6380370

GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: GTC-007

; CURRENT APPLICATION NUMBER: US/09/134,001C

; CURRENT FILING DATE: 1998-08-13

; PRIOR APPLICATION NUMBER: US 60/064,964

; PRIOR FILING DATE: 1997-11-08

; PRIOR APPLICATION NUMBER: US 60/055,779

; PRIOR FILING DATE: 1997-08-14

; NUMBER OF SEQ ID NOS: 5674

; SEQ ID NO 1597

; LENGTH: 927

; TYPE: DNA

; ORGANISM: Staphylococcus epidermidis

US-09-134-001C-1597

Query Match 5.8%; Score 18; DB 4; Length 927;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 217 AAAATGGAAGAAATTAGCT 234

Db 586 AAAATGGAAGAAATTAGCT 603

RESULT 2

US-09-480-921B-7/c
; Sequence 7, Application US/09480921B
; Patent No. 6387637

GENERAL INFORMATION:

; APPLICANT: Levin, Joshua Z.

; APPLICANT: Budziszewski, Gregory J.

; APPLICANT: Potter, Sharon L.

; APPLICANT: Wegrich, Lynette M.

; TITLE OF INVENTION: Herbicide Target Genes and Methods

; FILE REFERENCE: PB/5-30780A

; CURRENT APPLICATION NUMBER: US/09/480,921B

; CURRENT FILING DATE: 2000-01-11

; NUMBER OF SEQ ID NOS: 29

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 7

; LENGTH: 2163

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

```
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2163)
US-09-480-921B-7

Query Match          5.5%; Score 17; DB 4; Length 2163;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 275 AGAATCTACATCATCTT 291
Db 1872 AGAATCTACATCATCTT 1856

RESULT 3
US-09-480-921B-24/c
; Sequence 24, Application US/09480921B
; Patent No. 6387637
; GENERAL INFORMATION:
; APPLICANT: Levin, Joshua Z.
; APPLICANT: Budziszewski, Gregory J.
; APPLICANT: Potter, Sharon L.
; APPLICANT: Weirich, Lynette M.
; TITLE OF INVENTION: Herbicide Target Genes and Methods
; FILE REFERENCE: PB/3-30780A
; CURRENT APPLICATION NUMBER: US/09/480,921B
; CURRENT FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 24
; LENGTH: 5077
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-480-921B-24

Query Match          5.5%; Score 17; DB 4; Length 5077;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 275 AGAATCTACATCATCTT 291
Db 4287 AGAATCTACATCATCTT 4271

RESULT 4
US-08-651-155B-56
; Sequence 56, Application US/08651155B
; Patent No. 6365401
; GENERAL INFORMATION:
; APPLICANT: Mahan Dr., Michael J.
; APPLICANT: Conner Mr., Christopher P.
; APPLICANT: Hiethoff Mr., Douglas M.
; TITLE OF INVENTION: METHOD AND PROBES FOR THE IDENTIFICATION
; TITLE OF INVENTION: OF MICROBIAL GENES SPECIFICALLY INDUCED DURING HOST
; NUMBER OF SEQUENCES: 255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chrisman, Bynum & Johnson, P.C.
; STREET: 1900 Fifteenth Street
; CITY: Boulder
; STATE: CO
; COUNTRY: USA
; ZIP: 80302
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/651,155B
; FILING DATE: 17-MAY-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
```

```
; NAME: Petersen Mr., Steven C.
; REGISTRATION NUMBER: 36,238
; REFERENCE/DOCKET NUMBER: 17060.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 303/546-1300
; TELEFAX: 303/449-5426
; TELEX: ABA1475
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 423 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: YES
US-08-651-155B-56

Query Match          5.1%; Score 16; DB 4; Length 423;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 262 GCAGGGTAACAAAGA 277
Db 88 GCAGGGTAACAAAGA 103

RESULT 5
US-09-118-554-66/c
; Sequence 66, Application US/09118554A
; Patent No. 6365348
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOUNDS FOR DIAGNOSIS OF BREAST CANCER AND
; TITLE OF INVENTION: METHODS FOR THEIR USE
; FILE REFERENCE: 210121.450C1
; CURRENT APPLICATION NUMBER: US/09/118,554A
; CURRENT FILING DATE: 1998-07-17
; EARLIER APPLICATION NUMBER: 08/988,255
; EARLIER FILING DATE: 1997-12-24
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 66
; LENGTH: 831
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-118-554-66

Query Match          5.1%; Score 16; DB 4; Length 831;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 85 AGCAGAAGGAGCCCAA 100
Db 17 AGCAGAAGGAGCCCAA 2

RESULT 6
US-09-118-627-66/c
; Sequence 66, Application US/09118627A
; Patent No. 6379951
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF BREAST CANCER
; TITLE OF INVENTION: AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.446C1
; CURRENT APPLICATION NUMBER: US/09/118,627A
; CURRENT FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 66
```

LENGTH: 831
TYPE: DNA
ORGANISM: Homo sapien
US-09-118-627-66

Query Match 5.1%; Score 16; DB 4; Length 831;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 85 AGCAGAAGGCCCAA 100
DB 17 AGCAGAAGGCCCAA 2

RESULT 7

US-09-602-877A-66/c
Sequence 66, Application US/09602877A
Patent No. 6432707

GENERAL INFORMATION:

APPLICANT: Reed, Steven G.
APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Davin C.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
OF BREAST CANCER
FILE REFERENCE: 210121.446C5
CURRENT FILING DATE: 2000-06-22
NUMBER OF SEQ ID NOS: 107
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 66
LENGTH: 831
TYPE: DNA
ORGANISM: Homo sapien
US-09-602-877A-66

Query Match 5.1%; Score 16; DB 4; Length 831;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 85 AGCAGAAGGCCCAA 100
DB 17 AGCAGAAGGCCCAA 2

RESULT 8

US-09-889-595-2
Sequence 2, Application US/09889595
Patent No. 6410749

GENERAL INFORMATION:

APPLICANT: Aventis CropScience GmbH
TITLE OF INVENTION: PROMOTERS FOR GENE EXPRESSION IN CARYOPSES OF PLANTS
FILE REFERENCE: 514413-3895
CURRENT FILING DATE: 2001-07-05
PRIOR APPLICATION NUMBER: DE 100 32 379.0
PRIORITY FILING DATE: 2000-07-05
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 844
TYPE: DNA
ORGANISM: Triticum aestivum
US-09-889-595-2

Query Match 5.1%; Score 16; DB 4; Length 844;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 CAAGGAGATGTCTGG 32
DB 687 CAAGGAGATGTCTGG 702

RESULT 9

US-08-332-312-3/c
Sequence 3, Application US/08332312
Patent No. 5559026

GENERAL INFORMATION:

APPLICANT: PFICE, Laura A.
APPLICANT: Pausch, Mark H.
TITLE OF INVENTION: Functional Expression of a Drosophila
Melanogaster Putative Potassium Channel in Yeast
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: American Cyanamid Company
STREET: One Cyanamid Plaza
CITY: Wayne
STATE: New Jersey
COUNTRY: US
ZIP: 07470-8426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/332.312
FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Harrington, James J.
REGISTRATION NUMBER: P-38,711
REFERENCE/DOCKET NUMBER: 32,421
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-831-3246
TELEFAX: 201-831-3305

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:
LENGTH: 1011 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1008
US-08-332-312-3

Query Match

Best Local Similarity 5.1%; Score 16; DB 1; Length 1011;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 140 ACAATATACACCATAT 155

DB 187 ACAATATACACCATAT 172

RESULT-10

US-08-961-083-171

Sequence 171, Application US/08961083
Patent No. 6159469

GENERAL INFORMATION:

APPLICANT: Choi et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines
NUMBER OF SEQUENCES: 452
CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2

Issue date, Sept 24, 1996

SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,083
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 171:
SEQUENCE CHARACTERISTICS:
LENGTH: 1282 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-961-083-171

Query Match 5.1%; Score 16; DB 3; Length 1282;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 219 AATGGAAGAAATTAGCT 234
|||||
DB 445 AATGGAAGAAATTAGCT 460

RESULT 11
US-09-026-039-1
Sequence 1, Application US/09026039
Patent No. 6329567
GENERAL INFORMATION:
APPLICANT: Jofuku, K. Diane
APPLICANT: Okamoto, Jack K.
TITLE OF INVENTION: Methods for Improving Seeds
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/026,039
FILING DATE: 19-FEB-1998
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/912,272
FILING DATE: 15-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/879,827
FILING DATE: 20-JUN-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/700,152
FILING DATE: 20-AUG-1996
NAME: Bastian, Kevin L.
REGISTRATION NUMBER: 34,774
REFERENCE/DOCKET NUMBER: 023070-067230US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1669 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: -
LOCATION: 1..1669
OTHER INFORMATION: /note= "canola APETALA2 (AP2) domain
OTHER INFORMATION: containing (ADC) gene"
FEATURE:
NAME/KEY: misc feature
LOCATION: 1..532
OTHER INFORMATION: /note= "gene sequences 5' to start site"
US-09-026-039-1

Query Match 5.1%; Score 16; DB 4; Length 1669;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 278 ATCTACATCATCTTGC 293
|||||
DB 153 ATCTACATCATCTTGC 178

RESULT 12
US-08-811-897A-39
Sequence 39, Application US/08811897A
Patent No. 5858787
GENERAL INFORMATION:

APPLICANT: ONDA, Haruo
APPLICANT: OHTAKI, Tetsuya
APPLICANT: MASUDA, Yasushi
APPLICANT: KITADA, Chieko
APPLICANT: ISHIBASHI, Yoshihiro
APPLICANT: HOSOYA, Masaki
APPLICANT: OGI, Kazuhiro
APPLICANT: MIYAMOTO, Yasunori
APPLICANT: HABATA, Yugo
APPLICANT: SHIMAMOTO, No. 5858787Lo
TITLE OF INVENTION: PACAP RECEPTOR PROTEIN, METHOD FOR PREPARING
TITLE OF INVENTION: SAID PROTEIN, AND USE THEREOF
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
ADDRESSEE: CUSHMAN
STREET: 130 Water Street
CITY: Boston
STATE: Massachusetts
COUNTRY: US
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/811,897A
FILING DATE: 05-MAR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/202,986
FILING DATE: February 25, 1994
ATTORNEY/AGENT INFORMATION:
NAME: RESNICK, David S.
REGISTRATION NUMBER: 34235
REFERENCE/DOCKET NUMBER: 44168-DIV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)523-3400
TELEFAX: (617)523-6440
INFORMATION FOR SEQ ID NO: 39:

ISSUE Jan 12, 1999

SEQUENCE CHARACTERISTICS:
LENGTH: 2730 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cdna
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 498..1952
US-08-811-897A-39

Query Match 5.1%; Score 16; DB 2; Length 2730;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 71 CAGCTTGATGCCAG 86
DB 2465 CAGCTTGATGCCAG 2480

RESULT 13
US-08-855-213-39
Sequence 39, Application US/08855213
Patent No. 5892004
GENERAL INFORMATION:
APPLICANT: ONDA, Haruo
APPLICANT: OHTAKI, Tetsuya
APPLICANT: MASUDA, Yasushi
APPLICANT: KITADA, Chieko
APPLICANT: ISHIBASHI, Yoshihiro
APPLICANT: HOSOYA, Masaki
APPLICANT: OGI, Kazuhiro
APPLICANT: MIYAMOTO, Yasunori
APPLICANT: HABAYA, Yugo
APPLICANT: SHIMAMOTO, No. 5892004io
TITLE OF INVENTION: PACAP RECEPTER PROTEIN, METHOD FOR
PREPARING SAID PROTEIN, AND USE THEREOF
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
ADDRESSEE: CUSHMAN
STREET: 130 Water Street
CITY: Boston
STATE: Massachusetts
COUNTRY: US
ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/855,213
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/202,986
FILING DATE: 25-FEB-1994
ATTORNEY/AGENT INFORMATION:
NAME: RESNICK, David S.
REGISTRATION NUMBER: 34235
REFERENCE/DOCKET NUMBER: 44168
TELEPHONE: (617)523-3400
TELEFAX: (617)523-6440
TELEX: 200291 STRE UR
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 2730 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cdna

FEATURE:
NAME/KEY: mat_peptide
LOCATION: 498..1952
US-08-855-213-39
Query Match 5.1%; Score 16; DB 2; Length 2730;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 71 CAGCTTGATGCCAG 86
DB 2465 CAGCTTGATGCCAG 2480

RESULT 14
US-09-201-474-39
Sequence 39, Application US/09201474
Patent No. 6399316
GENERAL INFORMATION:
APPLICANT: ONDA, Haruo
APPLICANT: OHTAKI, Tetsuya
APPLICANT: MASUDA, Yasushi
APPLICANT: KITADA, Chieko
APPLICANT: ISHIBASHI, Yoshihiro
APPLICANT: HOSOYA, Masaki
APPLICANT: OGI, Kazuhiro
APPLICANT: MIYAMOTO, Yasunori
APPLICANT: HABAYA, Yugo
APPLICANT: SHIMAMOTO, No. 6399316io
TITLE OF INVENTION: PACAP RECEPTER PROTEIN, METHOD FOR PREPARING
SAID PROTEIN, AND USE THEREOF
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
ADDRESSEE: CUSHMAN
STREET: 130 Water Street
CITY: Boston
STATE: Massachusetts
COUNTRY: US
ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/201,474
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/811,897
FILING DATE: 05-MAR-1997
ATTORNEY/AGENT INFORMATION:
NAME: RESNICK, David S.
REGISTRATION NUMBER: 34235
REFERENCE/DOCKET NUMBER: 44168-DIV
TELEPHONE: (617)523-3400
TELEFAX: (617)523-6440
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 2730 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cdna
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 498..1952
US-09-201-474-39

Query Match 5.1%; Score 16; DB 4; Length 2730;
Best Local Similarity 100.0%; Pred. No. 39;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 71 CAGCTTGGATGGCCAG 86
|||||

Db 2465 CAGCTTGGATGGCCAG 2480

Search completed: July 7, 2003, 20:06:47
Job time : 66 secs

RESULT 15
US-08-811-897A-38
; Sequence 38, Application US/08811897A
; Patent No. 5858787
; GENERAL INFORMATION:
; APPLICANT: ONDA, Haruo
; APPLICANT: OHTAKI, Tetsuya
; APPLICANT: MASUDA, Yasushi
; APPLICANT: KITADA, Chieko
; APPLICANT: ISHIBASHI, Yoshihiro
; APPLICANT: HOSOYA, Masaki
; APPLICANT: OGI, Kazuhiro
; APPLICANT: MIYAMOTO, Yasunori
; APPLICANT: HABATA, Yugo
; APPLICANT: SHIMAMOTO, No. 585878710
; TITLE OF INVENTION: PACAP RECEPTOR PROTEIN, METHOD FOR PREPARING
; TITLE OF INVENTION: SAID PROTEIN, AND USE THEREOF
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
; ADDRESSEE: CUSHMAN
; STREET: 130 Water Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/811,897A
; FILING DATE: 05-MAR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/202,986
; FILING DATE: February 25, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: RESNICK, David S.
; REGISTRATION NUMBER: 34235
; REFERENCE/DOCKET NUMBER: 44168-DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)523-3400
; TELEFAX: (617)523-6440
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2814 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 498..2036
US-08-811-897A-38

Query Match 5.1%; Score 16; DB 2; Length 2814;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 71 CAGCTTGGATGGCCAG 86
|||||

Db 2549 CAGCTTGGATGGCCAG 2564

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 7, 2003, 19:14:44 ; Search time 1434 Seconds

(without alignments)

3512.410 Million cell updates/sec

Title: US-09-820-425B-12

Perfect score: 311

Sequence: 1 actccagctctgtgtgcaag.....gccatttgcaagtaagctt 311

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size : 0

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST:*

- 1: em_estba.*
- 2: em_esthum.*
- 3: em_estin.*
- 4: em_estum.*
- 5: em_estov.*
- 6: em_estpl.*
- 7: em_estro.*
- 8: em_htc.*
- 9: gb_est1.*
- 10: gb_est2.*
- 11: gb_htc.*
- 12: gb_est3.*
- 13: gb_est4.*
- 14: gb_est5.*
- 15: em_estfun.*
- 16: em_estom.*
- 17: gb_gss.*
- 18: em_gss_hum.*
- 19: em_gss_inv.*
- 20: em_gss_pln.*
- 21: em_gss_vrt.*
- 22: em_gss_fun.*
- 23: em_gss_mam.*
- 24: em_gss_mus.*
- 25: em_gss_other.*
- 26: em_gss_pro.*
- 27: em_gss_rod.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21	6.8	523	10	BB768985
c 2	21	6.8	562	14	BQ089932
3	20	6.4	682	13	BJ517473
c 4	20	6.4	694	17	AG088659
5	20	6.4	917	12	BG106177
c 6	20	6.4	1063	17	CNS06RYM

7	19	6.1	180	14	F03939	F03939 HSC2EB112 n
8	19	6.1	232	12	BF447312	BF447312 7p47f02.x
c 9	19	6.1	235	9	AI912227	AI912227 wd71q06.x
c 10	19	6.1	241	10	BB013037	BB013037 BB013037
c 11	19	6.1	256	12	BF359929	BF359929 CM2-MT009
c 12	19	6.1	264	9	AV278668	AV278668 AV278668
c 13	19	6.1	284	12	BG451432	BG451432 NF107D07D
c 14	19	6.1	294	12	BE812587	BE812587 QV1-AN005
15	19	6.1	299	9	AI700400	AI700400 wd07f08.x
16	19	6.1	307	9	AI698350	AI698350 tx63f05.x
c 17	19	6.1	310	9	AI697019	AI697019 wc85a02.x
c 18	19	6.1	323	10	BB558828	BB558828 BB558828
c 19	19	6.1	329	9	AA992991	AA992991 ct76e01.s
c 20	19	6.1	335	13	BI195977	BI195977 603756281
21	19	6.1	335	14	F02963	F02963 HSC1HH062 n
22	19	6.1	341	9	AI950680	AI950680 wx52h10.x
23	19	6.1	352	9	AI917570	AI917570 to27g07.x
c 24	19	6.1	352	14	BM738742	BM738742 K-EST0006
25	19	6.1	366	9	AI651053	AI651053 wa20g05.x
26	19	6.1	373	9	AI565458	AI565458 tq76c03.x
27	19	6.1	397	9	AA643721	AA643721 nl96g10.s
28	19	6.1	409	10	AM662614	AM662614 hi33d12.x
29	19	6.1	428	10	AM473443	AM473443 hb04h05.x
30	19	6.1	432	9	AI953258	AI953258 wq02d09.x
31	19	6.1	436	9	AI373323	AI373323 qz45b11.x
32	19	6.1	438	12	BF509344	BF509344 UI-H-B14-
33	19	6.1	448	14	H16805	H16805 vm39d07.s1
34	19	6.1	459	12	BF512779	BF512779 UI-H-BW1-
35	19	6.1	462	9	AI680075	AI680075 tw64g07.x
36	19	6.1	463	9	AI224511	AI224511 qi34g05.x
37	19	6.1	465	13	BJ386935	BJ386935 BJ386935
c 38	19	6.1	468	9	AI855811	AI855811 sc29f04.y
c 39	19	6.1	477	12	BF597597	BF597597 su98h10.y
c 40	19	6.1	483	10	AM054875	AM054875 ws96a06.x
c 41	19	6.1	499	10	AM455769	AM455769 FK9903181
c 42	19	6.1	501	13	BI747027	BI747027 rm34h12.y
c 43	19	6.1	518	13	BJ361124	BJ361124 BJ361124
44	19	6.1	524	9	AI913628	AI913628 tz90c06.x
c 45	19	6.1	529	13	BM088637	BM088637 502154 MA

ALIGNMENTS

RESULT 1	BB768985	523 bp	linear	EST 17-OCT-2001
LOCUS	BB768985	RIKEN full-length enriched, B16 F10Y cells	Mus musculus	
DEFINITION	CDNA clone G370090006 3', mRNA sequence.			
ACCESSION	BB768985			
VERSION	BB768985.1	GI:16211527		
KEYWORDS	EST.			
SOURCE	house mouse.			
ORGANISM	Mus musculus			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
AUTHORS	1 (bases 1 to 523)			
	Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K., Ishii,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,H., Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Watanishi,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.			
TITLE	RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al. 2001)			
JOURNAL	Unpublished (2001)			
COMMENT	Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan			

Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gsr.riken.go.jp,
 URL: http://genome.gsc.riken.go.jp/
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh
 M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
 Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura
 S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and
 Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.
 10 (11), 1757-1771 (2000)
 Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara
 Y. and Hayashizaki, Y.
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 Please visit our web site (http://genome.gsc.riken.go.jp) for
 further details.
 e mouse tissues.
 Location/Qualifiers
 1. .523
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="G370090006"
 /cell_lib="RIKEN full-length enriched, B16 F10Y cells"
 /note="pooled tissues; (tissue_type=cerebellum,
 dev_stage=16 days neonate, sex=mixed),
 (tissue_type=cerebellum, dev_stage=0 day neonate,
 sex=mixed), (tissue_type=hippocampus, dev_stage=adult,
 sex=male), (tissue_type=whole body, dev_stage=9 days
 embryo, sex=mixed), (tissue_type=lung, dev_stage=13 days
 embryo, sex=mixed)"
 150 a 124 c 97 g 152 t
 BASE COUNT
 ORIGIN
 Query Match 6.8%; Score 21; DB 10; Length 523;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 257 AAAAGCAGGTAACAAAGA 277
 Db 279 AAAAGCAGGTAACAAAGA 299
 RESULT 2
 BQ089932/c
 LOCUS
 DEFINITION
 rc54g02.y2 Meloidogyne hapla egg PAMPl v1 Meloidogyne hapla cDNA 5'
 similar to TR:Q19719 Q19719 COSMID F22D3. [1] ; mRNA sequence.
 ACCESSION
 BQ089932
 VERSION
 BQ089932.1 GI:20064133
 KEYWORDS
 EST.
 SOURCE
 Meloidogyne hapla.
 ORGANISM
 Meloidogyne hapla.
 Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchida;
 Tylenchoidea; Heteroderidae; Meloidogyninae; Meloidogyne.
 1 (bases 1 to 562)
 McCarter, J., Clifton, S., Chiapelli, B., Pape, D., Martin, J., Wylie, T.,
 Dante, M., Marra, M., Hillier, L., Kucaba, T., Theising, B., Bowers, Y.,
 Gibbons, M., Ritter, E., Bennett, J., Franklin, C., Tsagarelshvili, R.,
 Ronko, I., Kennedy, S., Maguire, L., Beck, C., Underwood, K., Steptoe
 M., Allen, M., Person, B., Swaller, T., Harvey, N., Schurk, R., Kohn, S.,
 Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and
 Wilson, R.
 The Washington Univ. Nematode EST Project, 1999
 Unpublished (1999)

COMMENT
 Contact: McCarter JP
 The Washington Univ. Nematode EST Project, 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 The library was constructed by Claire Murphy and Dr. James McCarter
 at Washington University, St. Louis. DNA Sequencing by: Washington
 University Genome Sequencing Center
 Possible reversed clone: similarity on wrong strand
 Seq primer: -40RP from Gibco
 High quality sequence stop: 411.
 Location/Qualifiers
 1. .562
 /organism="Meloidogyne hapla"
 /db_xref="taxon:6305"
 /clone_lib="Meloidogyne hapla egg PAMPl v1"
 /dev_stage="parasitic adult females"
 /lab_host="DH10b"
 /note="Vector: PAMPl (Gibco); Site1: NotI; Site2: SalI;
 The library was constructed by Claire Murphy and Dr. James
 McCarter at Washington University, St. Louis. The cDNA was
 made by using Dynabead oligo-dT priming (Dyna). PCR based
 library using a modified protocol from the SMART PCR cDNA
 Synthesis Kit from Clontech. Directionally cloned into the
 UDG sites of PAMPl."
 188 a 89 c 70 g 215 t
 BASE COUNT
 ORIGIN
 Query Match 6.8%; Score 21; DB 14; Length 562;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 215 AAAAATGGAAGTAATAGCTA 235
 Db 79 AAAAATGGAAGTAATAGCTA 59
 RESULT 3
 BJ517473
 LOCUS
 DEFINITION
 BJ517473 MF01FSA cDNA Oryzias latipes cDNA clone MF01FSA039L05 3'
 mRNA sequence.
 ACCESSION
 BJ517473
 VERSION
 BJ517473.1 GI:22176285
 KEYWORDS
 EST.
 SOURCE
 Japanese medaka.
 ORGANISM
 Oryzias latipes
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorpha; Acanthopterygii; Percormorpha; Atherinomorpha;
 Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
 1 (bases 1 to 682)
 Kohara, Y., Shin, I., Kimura, T., Narita, T., Jindo, T. and Takeda, H.
 Medaka EST project in Takeda's lab
 Unpublished (2001)
 Contact: Tadasu Shin-i
 Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tsunigenes.nig.ac.jp.
 Location/Qualifiers
 1. .682
 /organism="Oryzias latipes"
 /strain="d-r"
 /db_xref="taxon:8090"
 /clone="MF01FSA039L05"
 /clone_lib="MF01FSA cDNA"
 /sex="mixture of female and male"


```

/tissue_type="whole embryo"
/dev_stage="fry stage 40"
183 a 164 c 148 g 187 t

BASE COUNT      6.4%; Score 20; DB 13; Length 682;
ORIGIN           100.0%; Pred. No. 40;
Query Match     20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Best Local Similarity

205 CATTGAGTGAAGAAATGGA 224
|||||
12 CATTGAGTGAAGAAATGGA 31
Db

RESULT 4
AG088659/c
LOCUS
DEFINITION
Pan troglodytes DNA, clone: PTB-087K22.F, genomic survey sequence.
ACCESSION
AG088659.1 GI:16640461
VERSION
GSS.
KEYWORDS
SOURCE
Pan troglodytes male lymphoblast DNA, clone_lib:PTB Chimpanzee Male
BAC Library clone:PTB-087K22.F.
ORGANISM
Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
1
Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
BAC end sequences of Library PTB
Unpublished
2 (bases 1 to 694)
Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: chimpesgsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
Clones are derived from the chimpanzee BAC library PTB This BAC end
was generated during the RAD process and may have higher chance of
clone tracking errors.
PRIMERS
Sequencing: -21M13
LIBRARY
Vector : pKSI45
R.Site 1 : SacI
R.Site 2 : SacI.
Location/Qualifiers
1. .694
/organism="Pan troglodytes"
/db_xref="taxon:9598"
/clone="PTB-087K22.F"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"
226 a 137 c 104 g 227 t
BASE COUNT      6.4%; Score 20; DB 17; Length 694;
ORIGIN           100.0%; Pred. No. 40;
Query Match     20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Best Local Similarity

209 CAGTGAAGAAATGGAAGAA 228
|||||
565 CAGTGAAGAAATGGAAGAA 546
Db

RESULT 5
AG106177
LOCUS
DEFINITION
602289510F1 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:4384442 5',
EST 30-JAN-2001
BASE COUNT      6.4%; Score 20; DB 12; Length 917;
ORIGIN           100.0%; Pred. No. 41;
Query Match     20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Best Local Similarity

259 AAAGCAGGGTACAAAAGAA 278
|||||
853 AAAGCAGGGTACAAAAGAA 872
Db

RESULT 6
CNS06RVM/c
LOCUS
DEFINITION
T3 end of clone AW0AA018G04 of library AWOAA from strain CLIB 89 of
Yarrowia lipolytica, genomic survey sequence.
ACCESSION
AL412580
VERSION
AL412580.1 GI:12183421
KEYWORDS
GSS.
SOURCE
Yarrowia lipolytica.
ORGANISM
Yarrowia lipolytica
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Dipodascaceae; Yarrowia.
1 (bases 1 to 1063)
Soulciet,J.I., Aigle,M., Artiguenave,F., Blandin,G.,
Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,
de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Lorente,B.,
Malpertuy,A., Neuveglise,C., Ozier-Kalogeropoulos,O., Potier,S.,
Saurin,W., Tekala,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
Wincker,P. and Weissenbach,J.
Genomic exploration of the hemiascomycetous yeasts: 1. A set of
yeast species for molecular evolution studies
FEBS Lett. 487 (1), 3-12 (2000)
2 (bases 1 to 1063)
JOURNAL
MEDLINE
PUBMED
20584711
11152876
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
2 (bases 1 to 1063)

```

AUTHORS
 Casaregola, S., Neuveglise, C., Lepingle, A., Bon, E., Feynrol, C.,
 Artiguenave, F., Wincker, P. and Gaillardin, C.
TITLE
 Genomic exploration of the hemiascomycetous yeasts: 17. Yarrowia
 lipolytica
JOURNAL
 FEBS Lett. 487 (1), 95-100 (2000)
MEDLINE
 20584727
PUBMED
 11152892
REFERENCE
 3 (bases 1 to 1063)
AUTHORS
 Genoscope.
TITLE
 Direct Submission
JOURNAL
 Submitted (07-SEP-2000) Genoscope - Centre National de Sequencage,
 2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
 seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT
 This GSS is part of a random genomic sequencing program of thirteen
 yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
 exiguus, Saccharomyces kluyveri, Zygosaccharomyces rouxii,
 Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
 lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
 angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
 Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
 5 kb were prepared and both extremities were sequenced. See
 keywords for description of this sequence and for the sequence of
 the other extremity of this insert.

FEATURES
 source
 1..1063
 /organism="Yarrowia lipolytica"
 /strain="CLIB 89"
 /db_xref="taxon:4952"
 /clone="AW0A018G04"
 /clone_lib="AW0AA"
 /note="end : T3"
 BASE COUNT 289 a 287 c 220 g 263 t 4 others
 ORIGIN

Query Match 6.4%; Score 20; DB 17; Length 1063;
 Best Local Similarity 100.0%; Pred. NO. 41;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 157 TACTAGCTGTATGTTGT 176
 |||||
 Db 112 TACTAGCTGTATGTTGT 93

RESULT 7
 F03939
 LOCUS
 DEFINITION
 HSC2EB112 normalized infant brain cDNA Homo sapiens CDNA clone
 C-2eb11 3', mRNA sequence.
 F03939
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 180)
 Authors
 Auffray, C., Behar, G., Bois, F., Bouchier, C., da Silva, C., Devignes,
 M.D., Duprat, S., Houlgatte, R., Jumeau, M.N., Lamy, B., Lorenzo, F.,
 Mitchell, H., Mariage-Samson, R., Pietu, G., Pouliot, Y.,
 Sebastiani-Kabaktchis, C. and Tessier, A.
TITLE
 IMAGE: molecular integration of the analysis of the human genome
 and its expression
JOURNAL
 C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)
MEDLINE
 95277534
COMMENT
 Contact: Genethon
 Genethon
 Genethon Centre de recherche sur le Genome Humain
 1, rue de l'Internationale, BP60 91002 EVRY Cedex, FRANCE
 Tel: 33169472800
 Fax: 33160778698
 Email: genexpress@genethon.fr
 Single read, removed at sequence 5' end
 -genexpress_library_idt: C; Genexpress_sequence_idt: alc-2eb11

FEATURES
 source
 1..180
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="c-2eb11"
 /clone_lib="normalized infant brain cDNA"
 /sex="Female"
 /tissue_type="total brain"
 /dev_stage="3 months old"
 /note="Organ: brain; Vector: lafmid BA; Site: 1: HindIII;
 Site 2: NotI; sex: Female; dev_stage=3 months old;
 isolate-muscular atrophy patient; tissue_type=total brain
 ; total mRNA was oligo-(dT) primed and directionally
 cloned 5' -> 3' into the HindIII -> NotI sites of the
 lafmid BA vector. Clone library from B. Soares, Psychiatry
 Dept. Columbia University, USA. Normalization_method:
 Bento Soares, P.N.A.S. in press"
 BASE COUNT 52 a 53 c 24 g 49 t 2 others
 ORIGIN

Query Match 6.1%; Score 19; DB 14; Length 180;
 Best Local Similarity 100.0%; Pred. NO. 1.2e+02;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 142 AATATACACCATATATACT 160
 |||||
 Db 76 AATATACACCATATATACT 94

RESULT 8
 BF447312
 LOCUS
 DEFINITION
 232 bp mRNA linear EST 01-DEC-2000
 7p47f02.x1 NCI CGAP Pr28 Homo sapiens CDNA clone IMAGE:3648915 3'
 similar to contains element L1 repetitive element ;, mRNA sequence.
 BF447312
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 232)
 Authors
 NCI CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
JOURNAL
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-re@mail.nih.gov
 Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: M. Bento Soares, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL, send email to:
 info@image.llnl.gov
 Seq primer: -40UP from Gibco.

FEATURES
 source
 1..232
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3648915"
 /clone_lib="NCI CGAP Pr28"
 /sex="male"
 /dev_stage="adult"
 /lab_host="DH10B"
 /note="Organ: prostate; Vector: pTT3D-Pac (Pharmacia)
 with a modified polylinker; Plasmid DNA from the
 normalized library NCI CGAP Pr22 was prepared, and ss
 circles were made in vitro. Following HAP purification,
 this DNA was used as tracer in a subtractive hybridization
 reaction. The driver was PCR-amplified cDNAs from a pool

QY	142	AATATACACCATATATACT	160
Dd	176	AATATACACCATATATACT	198

FEATURES

```

Location/Qualifiers
1. .241
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="4930458L07"
/clone_lib="RIKEN full-length enriched, adult male testis
(DH10B)"
/sex="male"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/note="Site_1: Salt; Site_2: BamHI; CDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN, Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'-
GAGAGAGAGAGAGATCCAGAGCTCTTTTTTTTTTTTTTTN 3'], cDNA was
prepared by using trihalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by

```

cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGATTCTCAGTTAATAATTAATCCCTCCCTCCCTCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI."

BASE COUNT 78 a 43 c 28 g 92 t
 ORIGIN
 Query Match 6.1%; Score 19; DB 10; Length 241;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 204 CCATTCAGTGAATAATG 222
 Db 177 CCATTCAGTGAATAATG 159

RESULT 11
 BF359929
 LOCUS 256 bp mRNA linear EST 22-NOV-2000
 DEFINITION CM2-MT0099-190700-285-C07 MT0099 Homo sapiens cDNA, mRNA sequence.
 ACCESSION BF359929
 VERSION BF359929.1 GI:11319001
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo. 1 (bases 1 to 256)
 AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M. R., Nagai, M. A., da Silva, W. Jr., Zago, M. A., Bordin, S., Costa, F. F., Goldman, G. H., Carvalho, A. F., Matsukuma, A., Baia, G. S., Simpson, D. H., Brunstein, A., de Oliveira, P. S., Bucher, P., Jongeneel, C. V., O'Hare, M. J., Soares, F., Brentani, R. R., Reis, L. F., de Souza, S. J., and Simpson, A. J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM2&t2=CM2-MT0099-190700-285-C07&t3=2000-07-19&t4=1)
 Seq primer: puc 18 forward
 High quality sequence stop: 256.
 FEATURES
 Location/Qualifiers
 1..256
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="MT0099"
 /dev_stage="Adult"
 /note="Organ: marrow; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196 716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 74 a 59 c 42 g 81 t
 ORIGIN
 Query Match 6.1%; Score 19; DB 12; Length 256;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 212 TGAATAAATGGAAGAATT 230
 Db 56 TGAATAAATGGAAGAATT 74

RESULT 12
 AV278668/c
 LOCUS 264 bp mRNA linear EST 05-NOV-1999
 DEFINITION AV278668 RIKEN full-length enriched, adult male testis (DH10B) Mus musculus cDNA clone 4933403C22 3', mRNA sequence.

ACCESSION AV278668
 VERSION AV278668.1 GI:6266705
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 264)

AUTHORS Konno, H., Aizawa, K., Akahira, S., Akiyama, J., Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirozane, T., Hori, F., Ishii, Y., Ishikawa, T., Itoh, M., Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N., Kojima, Y., Koya, S., Kusakabe, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y., Owa, C., Ozawa, Y., Saito, H., Sano, M., Sato, K., Shibata, K., Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Takahashi, F., Tateo, M., Tominaga, N., Tsunoda, Y., Watahiki, A., Watanabe, S., Yamamura, T., Yasunishi, A., Yokota, T., Yoshiki, A., Yoshino, M., Muramatsu, M., and Hayashizaki, Y.
 RIKEN Mouse ESTs (Konno, H., et al. 1999)
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gsr.riken.go.jp,
 URL: http://genome.gsc.riken.go.jp/
 Sasakhi, N., Izawa, M., Watahiki, M., Ozawa, K., Tanaka, T., Yoneda, Y., Matsura, S., Carninci, P., Muramatsu, M., Okazaki, Y., and Hayashizaki, Y.

TITLE JOURNAL
 COMMENT

Transcriptional sequencing: A method for DNA sequencing using RNA polymerase. Proc. Natl. Acad. Sci. U.S.A. 95 (7), 3455-3460 (1998)
 Itoh, M., Kitzunai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J., Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y., and Hayashizaki, Y.
 Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)
 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
 Please visit our web site (http://genome.rtc.riken.go.jp) for further details.

FEATURES
 Location/Qualifiers
 1..264
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="4933403C22"
 /clone_lib="RIKEN full-length enriched, adult male testis (DH10B)"
 /sex="male"
 /tissue_type="testis"
 /dev_stage="adult"
 /lab_host="DH10B"
 /note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was

84 a 52 c 37 q 91 t

191 CCATTCACTGAAAAAATG 173

SOURCE	ORGANISM	Medicatio trunc
...

Plant Biology Di

Fax: 580 221 7380

Insert Length: 284 Std Error: 0.00
plate: 107 row: 5 column: 07

FEATURES

```
/organism="Medicago truncatula"
```

```
/clone NF10/D0/D1
/clone Lib="Drought"
```

```
/dev_stage="Pooled timepoints"
```

BASE COUNT	105 a	46 c	39 σ	94 +
timepoints:				

Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 19; Conservative 0; Mismatches 0; Traded

QY 214 AAAAAATGAAGAATTAG 232

Db 174 AAAAAATGGAAGATTAG 156

REVISION
VERSION

EST.
EST.

SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 299)
 REFERENCE NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 TITLE Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaps-r@mail.nih.gov
 Tissue Procurement: Elias Campo, M.D., Michael R. Emmert-Buck, M.D.,
 Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D.
 CDNA Library Arraying: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CCAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
 Insert Length: 379 Std Error: 0.00
 Seq primer: -40UP from Gibco.
 FEATURES Location/Qualifiers
 1..299
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2327439"
 /clone_lib="NCI-CCAP_Co3"
 /sex="pooled"
 /tissue_type="colon"
 /lab_host="DH10B"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
 was prepared from 12 pooled bulk tumor samples and primed
 with a Not I - oligo(dT) primer. Double-stranded cDNA was
 ligated to Eco RI adaptors (Pharmacia), digested with Not
 I and cloned into the Not I and Eco RI sites of the
 modified pT73 vector. Library went through one round of
 normalization."
 BASE COUNT 84 a 84 c 48 g 83 t
 ORIGIN

Query Match 6.1%; Score 19; DB 9; Length 299;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 142 AATATACACCATATATACT 160
 |||||
 Db 35 AATATACACCATATATACT 53

Search completed: July 7, 2003, 20:05:32
 Job time : 1439 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 7, 2003, 18:31:30 ; Search time 219 Seconds
(without alignments)
3198.043 Million cell updates/sec

Title: US-09-820-425B-12
Perfect score: 311
Sequence: 1 actcagctctgtgtgcaag.....gccattgcaggtaaagctt 311

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2185239 seqs, 1125999159 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_101002.*

1: /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	311	100.0	311	AA19594	Human Lung cancer
2	29	9.3	29	AA19601	Human LSG Sgling020
3	22	7.1	22	AA19599	Human LSG Sgling020
c 4	21	6.8	7110	AAS46489	Tumour suppressor
c 5	21	6.8	7110	ABL70587	Chemically treated
c 6	21	6.8	7110	ABL33556	Human immune syste
c 7	21	6.8	7110	AAS61290	Human gene regulat
c 8	21	6.8	34688	ABQ67059	Human angiogenesis
c 9	21	6.8	2365589	ABA90521	Genomic sequence o

c 10	19	6.1	208	AAH87073	Human single nucle
c 11	19	6.1	307	AAC07206	Human secreted pro
c 12	19	6.1	562	AAC99087	Human cancer cell
c 13	19	6.1	587	AAH10915	Human cDNA clone (
c 14	19	6.1	1164	AAZ98048	Human secreted pro
c 15	19	6.1	1164	AAZ11661	Human secreted pro
c 16	19	6.1	1164	ABK69757	Human secreted pro
c 17	19	6.1	2708	ABK69862	Human secreted pro
c 18	19	6.1	3019	AAH14240	Human cDNA sequenc
c 19	19	6.1	143899	AAH38336	Genomic sequence e
c 20	18	5.8	18	AAZ19600	Human LSG Sgling020
c 21	18	5.8	662	ABV46453	Human prostate exp
c 22	18	5.8	664	ABV16655	Human prostate exp
c 23	18	5.8	759	AAH53277	S. epidermidis ope
c 24	18	5.8	927	AAH92134	Staphylococcus epi
c 25	18	5.8	1465	ABN52173	Human foetal liver
c 26	18	5.8	1465	ABAS3457	Human bone marrow
c 27	18	5.8	1465	AAK27180	Probe #1699 for ge
c 28	18	5.8	1465	AAI11766	Probe #1762 used t
c 29	18	5.8	1465	AAI33076	Probe #1686 used t
c 30	18	5.8	1465	AAI01695	Human genome-deriv
c 31	18	5.8	1521	ABSO1730	Human foetal liver
c 32	18	5.8	1521	ABA66047	Human bone marrow
c 33	18	5.8	1521	AAK40210	Probe #10911 for g
c 34	18	5.8	1521	AAI20978	Probe #14911 used
c 35	18	5.8	1521	AAI46225	Probe #6682 used t
c 36	18	5.8	1521	AAI06691	Human genome-deriv
c 37	18	5.8	2832	ABSI4258	Human genome-deriv
c 38	18	5.8	3005	ABL61894	Colon adenocarcino
c 39	18	5.8	3218	AAH54088	S. epidermidis gen
c 40	18	5.8	10250	ABV25411	Human prostate exp
c 41	18	5.8	10250	ABN80082	Human chemically m
c 42	18	5.8	12212	ABK12920	Human SLC26A2 (sol
c 43	18	5.8	12212	ABK12975	Human SLC26A2 (sol
c 44	18	5.8	32192	ABA15328	Human nervous syst
c 45	18	5.8	32192	AAZ05257	Human reproductive
c 46	18	5.8	32192	AAZ05532	Human reproductive

ALIGNMENTS

RESULT 1

AA19594

ID AA19594 standard; cDNA; 311 BP.

XX AA19594;

AC AA19594;

XX AA19594;

DT 18-DEC-2001 (first entry)

XX Human Lung cancer Specific Gene (LSG) Sgling020 cDNA.

DE Human Lung cancer Specific Gene; LSG; cytostatic; vaccine; gene therapy;

XX Human; Lung cancer Specific Gene; LSG; cytostatic; vaccine; gene therapy;

KW cancer; immune response; cell disorder; tissue; ss.

OS Homo sapiens.

XX WO200172775-A2.

PN WO200172775-A2.

XX 04-OCT-2001.

XX 29-MAR-2001; 2001WO-US10051.

PF 29-MAR-2000; 2000US-192921P.

XX (DIAD-) DIADEXUS INC.

XX Chen S, Hu P, Recipon H, Macina RA;

XX WPI; 2001-616467/71.

XX New lung cancer specific genes and polypeptides encoded by the genes,

PT useful for detecting, diagnosing, monitoring, staging, prognosticating,

PT imaging and treating lung cancer

```
XX PS Claim 1; Page 74; 76pp; English.
XX CC The present sequence is lung cancer specific genes (LSG) cDNA. LSG and
XX CC its protein are useful for diagnosing, detecting, monitoring, staging,
XX CC prognosticating, imaging and treating lung cancer in a patient by
XX CC determining the levels of LSG in cells, tissues or bodily fluids in a
XX CC patient and comparing the determined levels of LSG with levels of LSG
XX CC in cells, tissues or bodily fluids from a normal human control, where a
XX CC change in determined levels of LSG in the patient versus normal control
XX CC is associated with the presence of lung cancer. LSG is used for
XX CC identifying potential therapeutic agents for use in imaging and treating
XX CC lung cancer. LSG antibody labelled with paramagnetic ions or a
XX CC radioisotope is useful for imaging lung cancer in a patient. LSG antibody
XX CC conjugated to a cytotoxic agent is useful for treating lung cancer in a
XX CC patient. LSG vaccine is useful for inducing an immune response against a
XX CC LSG protein and for treating lung cancer in a patient. LSG and its
XX CC protein are useful as diagnostic markers for lung cancer and for
XX CC diagnosis and treatment of disorders of cells, tissues and organisms.
XX SQ Sequence 311 BP; 107 A; 54 C; 64 G; 86 T; 0 other;

Query Match          100.0%; Score 311; DB 22; Length 311;
Best Local Similarity 100.0%; Pred. No. 1.8e-149;
Matches 311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTCCAGCTCTGTGTCAGGAGCATGTGCTGGAAATGTCACAGCATCGTATAGCAAGAGC 60
DB 1 ACTCCAGCTCTGTGTCAGGAGCATGTGCTGGAAATGTCACAGCATCGTATAGCAAGAGC 60

QY 61 ATATTGGCAACAGCTTGGATGGCCAGCAGAGAGGCCCAATGTGTGATTCAATTCAC 120
DB 61 ATATTGGCAACAGCTTGGATGGCCAGCAGAGAGGCCCAATGTGTGATTCAATTCAC 120

QY 121 AGTCGAATTAATTAATCAATACATATACACCATATATAGTGTGTGTTCTA 180
DB 121 AGTCGAATTAATTAATCAATACATATACACCATATATAGTGTGTGTTCTA 180

QY 181 TACTATAGTGATTGACTTGAACTCCATTCAGTCAAAAAAATGGAAGAAATAGCTATTGT 240
DB 181 TACTATAGTGATTGACTTGAACTCCATTCAGTCAAAAAAATGGAAGAAATAGCTATTGT 240

QY 241 ATCCATATGGGATACAAAAAGCAGGGTAAACAAAGAAATCTACATCATCTTGCATTGC 300
DB 241 ATCCATATGGGATACAAAAAGCAGGGTAAACAAAGAAATCTACATCATCTTGCATTGC 300

QY 301 AGGTAAGCTT 311
DB 301 AGGTAAGCTT 311

RESULT 2
AAD19601
ID AAD19601 standard; DNA; 29 BP.
XX AC AAD19601;
XX XX
XX XX 18-DEC-2001 (first entry)
XX DE Human LSG Sglnq020 cDNA amplifying Q-PCR probe.
XX XX
XX KW Human; Lung cancer Specific Gene; LSG; cytostatic; vaccine; gene therapy;
XX KW cancer; immune response; cell disorder; tissue; PCR probe; ss.
XX OS Homo sapiens.
XX XX
XX PN WO200172775-A2.
XX PD 04-OCT-2001.
XX XX
XX PF 29-MAR-2001; 2001WO-US10051.
XX XX
XX PR 29-MAR-2000; 2000US-192921P.
XX XX
XX PA (DIAD-) DIADEXUS INC.
XX PI Chen S, Hu P, Recipon H, Macina RA;
XX DR WPI; 2001-616467/71.
XX XX
XX PT New lung cancer specific genes and polypeptides encoded by the genes,
XX PT useful for detecting, diagnosing, monitoring, staging, prognosticating,
```

```
XX PA (DIAD-) DIADEXUS INC.
XX PI Chen S, Hu P, Recipon H, Macina RA;
XX DR WPI; 2001-616467/71.
XX XX
XX PT New lung cancer specific genes and polypeptides encoded by the genes,
XX PT useful for detecting, diagnosing, monitoring, staging, prognosticating,
XX PT imaging and treating lung cancer
XX XX
XX PS Example 5; Page 59; 76pp; English.
XX CC The present sequence is a PCR probe used for amplifying lung cancer
XX CC specific genes (LSG) cDNA. LSG and its protein are useful for diagnosing,
XX CC detecting, monitoring, staging, prognosticating, imaging and treating
XX CC lung cancer in a patient by determining the levels of LSG in cells,
XX CC tissues or bodily fluids in a patient and comparing the determined levels
XX CC of LSG with levels of LSG in cells, tissues or bodily fluids from a
XX CC normal human control, where a change in determined levels of LSG in the
XX CC patient versus normal control is associated with the presence of lung
XX CC cancer. LSG is used for identifying potential therapeutic agents for use
XX CC in imaging and treating lung cancer. LSG antibody labelled with
XX CC paramagnetic ions or a radioisotope is useful for imaging lung cancer in
XX CC a patient. LSG antibody conjugated to a cytotoxic agent is useful for
XX CC treating lung cancer in a patient. LSG vaccine is useful for inducing an
XX CC immune response against a LSG protein and for treating lung cancer in a
XX CC patient. LSG and its protein are useful as diagnostic markers for lung
XX CC cancer and for diagnosis and treatment of disorders of cells, tissues
XX CC and organisms.
XX SQ Sequence 29 BP; 11 A; 5 C; 7 G; 6 T; 0 other;

Query Match          9.3%; Score 29; DB 22; Length 29;
Best Local Similarity 100.0%; Pred. No. 6.5e-05;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 AGCATCGTATAGCAAGAGCATATTGGCA 69
DB 1 AGCATCGTATAGCAAGAGCATATTGGCA 29

RESULT 3
AAD19599
ID AAD19599 standard; DNA; 22 BP.
XX AC AAD19599;
XX XX
XX XX 18-DEC-2001 (first entry)
XX DE Human LSG Sglnq020 cDNA amplifying Lng146 forward PCR primer.
XX XX
XX KW Human; Lung cancer Specific Gene; LSG; cytostatic; vaccine; gene therapy;
XX KW cancer; immune response; cell disorder; tissue; PCR primer; ss.
XX OS Homo sapiens.
XX XX
XX PN WO200172775-A2.
XX PD 04-OCT-2001.
XX XX
XX PF 29-MAR-2001; 2001WO-US10051.
XX XX
XX PR 29-MAR-2000; 2000US-192921P.
XX XX
XX PA (DIAD-) DIADEXUS INC.
XX PI Chen S, Hu P, Recipon H, Macina RA;
XX DR WPI; 2001-616467/71.
XX XX
XX PT New lung cancer specific genes and polypeptides encoded by the genes,
XX PT useful for detecting, diagnosing, monitoring, staging, prognosticating,
```


PT imaging and treating lung cancer -
 XX Example 5; Page 59; 76pp; English.
 XX
 CC The present sequence is a PCR primer used for amplifying lung cancer
 CC specific genes (LSG) cDNA. LSG and its protein are useful for diagnosing,
 CC detecting, monitoring, staging, prognosticating, imaging and treating
 CC lung cancer in a patient by determining the levels of LSG in cells,
 CC tissues or bodily fluids in a patient and comparing the determined levels
 CC of LSG with levels of LSG in cells, tissues or bodily fluids from a
 CC normal human control, where a change in determined levels of LSG in the
 CC patient versus normal control is associated with the presence of lung
 CC cancer. LSG is used for identifying potential therapeutic agents for use
 CC in imaging and treating lung cancer. LSG antibody labelled with
 CC paramagnetic ions or a radioisotope is useful for imaging lung cancer in
 CC a patient. LSG antibody conjugated to a cytotoxic agent is useful for
 CC treating lung cancer in a patient. LSG vaccine is useful for inducing an
 CC immune response against a LSG protein and for treating lung cancer in a
 CC patient. LSG and its protein are useful as diagnostic markers for lung
 CC cancer and for diagnosis and treatment of disorders of cells, tissues
 CC and organisms.
 XX
 CC Sequence 22 BP; 6 A; 3 C; 8 G; 5 T; 0 other;
 SQ
 Query Match 7.1%; Score 22; DB 22; Length 22;
 Best Local Similarity 100.0%; Pred. No. 0.25;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 17 CAAGGAGATGCTGGAATGTC 38
 Db 1 CAAGGAGATGCTGGAATGTC 22
 RESULT 4
 AAS46489/c
 ID AAS46489 standard; DNA; 7110 BP.
 XX
 AC AAS46489;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Tumour suppressor gene derived chemically modified sequence #211.
 XX
 KW Human; tumour suppressor gene; oncogene; antitumour; cytostatic;
 KW cancer; tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;
 KW cytosine methylation; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200168912-A2.
 XX
 PD 20-SEP-2001.
 XX
 PF 15-MAR-2001; 2001WO-EP02955.
 XX
 PR 15-MAR-2000; 2000DE-1013847.
 PR 06-APR-2000; 2000DE-1019058.
 PR 07-APR-2000; 2000DE-1019173.
 PR 30-JUN-2000; 2000DE-1032529.
 PR 01-SEP-2000; 2000DE-1043826.
 XX
 PA (EPIG-) EPIGENOMICS AG.
 XX
 PI Olek A, Piepenbrock C, Berlin K;
 XX
 DR WPI; 2001-602752/68.
 XX
 PT Fragments of chemically modified genes associated with tumour suppressor
 PT genes and oncogenes, useful in designing primers and probes for
 PT analysing diseases associated with cytosine methylation state e.g.
 PT cancer -
 PS Claim 1; SEQ ID No 211; 27pp; English.
 XX

XX The invention relates to a nucleic acid comprising a sequence of 18
 CC bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with
 CC bisulphite, of genes associated with tumour suppression and
 CC oncogenes having a sequence taken from 536 (actually 533 since
 CC numbers 408, 458 and 500 are missing from the sequence listing) sequences
 CC (SS) and sequences complementary to (SS). The nucleic acid may be a
 CC peptide nucleic acid-oligomer (PNA) of at least 9 nucleotides and may
 CC form part of a set of probes for detecting the cytosine methylation state
 CC and/or single nucleotide polymorphisms and also to be used in an
 CC array for analysing diseases associated with CpG dinucleotides e.g.
 CC cancers and tumours. The probes can also be used in a method for
 CC ascertaining genetic and/or epigenetic parameters for the diagnosis
 CC and/or therapy of existing diseases or the predisposition to specific
 CC diseases, by analysing cytosine methylations. The parameters may be
 CC compared to another set of genetic and/or epigenetic parameters, the
 CC differences serving as basis for diagnosis and/or prognosis events which
 CC are disadvantageous to patients. The present sequence is one of the
 CC 533 genomic sequences derived from tumour suppressor genes and
 CC oncogenes.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 7110 BP; 1749 A; 289 C; 1822 G; 3250 T; 0 other;
 Query Match 6.8%; Score 21; DB 22; Length 7110;
 Best Local Similarity 100.0%; Pred. No. 0.86;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 141 CAATATACACCATATATACTA 161
 Db 6679 CAATATACACCATATATACTA 6659
 RESULT 5
 ABL70587/c
 ID ABL70587 standard; DNA; 7110 BP.
 XX
 AC ABL70587;
 XX
 DT 01-JUL-2002 (first entry)
 XX
 DE Chemically treated cell signalling DNA sequence#239.
 XX
 KW Cell signalling; cytosine methylation; cell signalling disease;
 KW cancer; tumour; cytostatic; ds.
 XX
 OS Unidentified.
 XX
 PN WO200202807-A2.
 XX
 PD 10-JAN-2002.
 XX
 PF 29-JUN-2001; 2001WO-EP07471.
 XX
 PR 30-JUN-2000; 2000DE-1032529.
 PR 01-SEP-2000; 2000DE-1043826.
 XX
 PA (EPIG-) EPIGENOMICS AG.
 XX
 PI Olek A, Piepenbrock C, Berlin K;
 XX
 DR WPI; 2002-154758/20.
 XX
 PT Nucleic acid, useful for diagnosis and therapy of diseases associated
 PT with cell signalling e.g. cancer, comprises chemically modified genomic
 PT sequences of genes associated with cell signalling -
 XX
 PS Claim 1; SEQ ID NO 477; 24pp+sequence listing; English.
 XX
 CC The invention relates to a nucleic acid comprising a sequence of at least

CC 18 bases of a segment of chemically pretreated DNA of genes associated
 CC with cell signalling. The activity of the modified sequences of the
 CC invention may be described as cytotatic. The object of the invention is
 CC to provide the chemically modified DNA of genes associated with cell
 CC signalling, as well as oligonucleotides and/or PNA-oligomers for
 CC detecting cytosine methylations, as well as a method which is
 CC particularly suitable for the diagnosis and/or therapy of genetic and
 CC epigenetic parameters of genes associated with cell signalling. The
 CC chemically modified DNA provided by the invention is useful for diagnosis
 CC and therapy of diseases such as solid tumours and cancer. The sequences
 CC given in records ABL70111-ABL70626 represent chemically pre-treated
 CC genomic DNA's of genes associated with cell signalling.
 CC Note: The sequence data for this patent is not represented in the printed
 CC specification, but is based on sequence information supplied by the
 CC European Patent Office.

XX SQ Sequence 7110 BP; 1749 A; 289 C; 1822 G; 3250 T; 0 other;
 Query Match 6.8%; Score 21; DB 24; Length 7110;
 Best Local Similarity 100.0%; Pred. No. 0.86;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 141 CAATATACCATATATATACTA 161
 Db 6679 CAATATACCATATATATACTA 6659
 |||||

RESULT 6
 ABL33556/C
 ID ABL33556 standard; DNA; 7110 BP.

XX AC ABL33556;
 XX 26-MAR-2002 (first entry)
 XX Human immune system associated gene SEQ ID NO: 1529.

XX Human; immune system disease; cytosine methylation; antiasthmatic;
 KW antiarteriosclerotic; antihaemic; cytotatic; nootropic;
 KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
 KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
 KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
 KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
 KW gene; ds.

XX Homo sapiens.

XX WO200200928-A2.

XX 03-JAN-2002.

XX 02-JUL-2001; 2001WO-EP07537.

XX 30-JUN-2000; 2000DE-1032529.

XX 01-SEP-2000; 2000DE-1043826.

XX (EPIG-) EPIGENOMICS AG.

XX Olek A, Piepenbrock C, Berlin K;

XX WPI; 2002-130909/17.

XX Nucleic acid comprising fragment of chemically modified gene, useful
 PT for diagnosis and treatment of diseases associated with abnormal
 PT cytosine methylation

XX Claim 1; SEQ ID NO 1529; 32pp + Sequence Listing; German.

XX The present invention provides a number of human immune system associated
 CC genes which are modified by the methylation of cytosines. The sequences
 CC can be used in the diagnosis and treatment of immune system disorders,
 CC including eye diseases such as retinopathy, neovascular glaucoma and

CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
 CC diseases. The present sequence is a gene of the invention.

XX SQ Sequence 7110 BP; 1749 A; 289 C; 1822 G; 3250 T; 0 other;

Query Match 6.8%; Score 21; DB 24; Length 7110;
 Best Local Similarity 100.0%; Pred. No. 0.86;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 141 CAATATACCATATATATACTA 161
 Db 6679 CAATATACCATATATATACTA 6659
 |||||

RESULT 7
 AAS61290/C

ID AAS61290 standard; DNA; 7110 BP.

XX AC AAS61290;

XX 29-JAN-2002 (first entry)

XX Human gene regulation-associated gene oligonucleotide #245.

XX Human; Gene regulation-associated gene; severe combined immunodeficiency;
 KW cardiac damage; inflammatory response; Haemophilia; Werner syndrome;
 KW asthma; HDR syndrome; congenital heart defect; Saethre-Chotzen syndrome;
 KW renal disease; Preeclampsia; cardiac allograft vascular disease;
 KW colorectal cancer; thyroid cancer; oesophageal cancer; ds; tumour;
 KW immunostimulant; cardiant; antiinflammatory; coagulant; antiasthmatic;
 KW nephrotropic; gynecological; anti-tumour; immunosuppressive; cytostatic.

XX Homo sapiens.

XX WO200177375-A2.

XX 18-OCT-2001.

XX 06-APR-2001; 2001WO-EP03968.

XX 06-APR-2000; 2000DE-1019058.

XX 07-APR-2000; 2000DE-1019173.

XX 30-JUN-2000; 2000DE-1032529.

XX 01-SEP-2000; 2000DE-1043826.

XX (EPIG-) EPIGENOMICS AG.

XX Olek A, Piepenbrock C, Berlin K;

XX WPI; 2002-017470/02.

XX New nucleic acid sequences from chemically modified genes associated
 PT with gene regulation, useful for analysing cytosine methylations for
 PT diagnosis and therapy of diseases e.g. severe combined immunodeficiency
 PT disease

XX Disclosure; SEQ ID NO 251; 26pp; English.

XX The invention relates to 224 nucleic acid sequences comprising at least
 CC 18 bases of a chemically pretreated gene associated with gene regulation
 CC selected from 43 known genes (or complementary sequences). The
 CC chemical pretreatment converts cytosine bases unmethylated at the
 CC 5-position to uracil or another base with hybridisation behaviour
 CC dissimilar to cytosine, to enable analysis of cytosine methylations.
 CC The DNA sequences, oligomers (or sets/arrays) and method are
 CC useful in the diagnosis of diseases (or predisposition to diseases)
 CC associated with gene regulation and in therapy of such diseases, by
 CC enabling analysis of the cytosine methylation patterns of such genes,
 CC kits are provided. They are especially useful in diagnosis
 CC and therapy of e.g. severe combined immunodeficiency disease, cardiac
 CC disorders, haemophilia, solid tumours and cancer, Werner syndrome,

CC asthma, HDR syndrome, Saethre-Chotzen syndrome, renal disease,
CC preclampsia, graft versus-host disease. The present sequence is a
CC sequence included in the sequence data for this specification and is
CC associated with the human gene regulation-associated genes.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX
SQ Sequence 7110 BP; 1749 A; 289 C; 1822 G; 3250 T; 0 other;

Query Match 6.8%; Score 21; DB 24; Length 7110;
Best Local Similarity 100.0%; Pred. No. 0.86;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 141 CAATATACACCATATATACTA 161
Db 6679 CAATATACACCATATATACTA 6659
|||||

RESULT 8
ABQ67059/c
ID ID ABQ67059 standard; DNA; 34688 BP.
XX
XX AC ABQ67059;
XX
XX DT
XX
XX 28-AUG-2002 (first entry)
XX Human angiogenesis associated polynucleotide SEQ ID NO 89.
XX
XX Human; angiogenesis; methylation; eye disease; glaucoma; tumour;
KW Inflammation; rheumatoid arthritis; diabetic retinopathy; antiulcers;
KW macular degeneration; inflammatory bowel disease; Crohn's disease;
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW antiarteriosclerotic; ds.
XX
XX Homo sapiens.
OS
XX
XX WO200246454-A2.
XX
XX 13-JUN-2002.
XX
XX 06-DEC-2001; 2001WO-EPI4320.
XX
XX 06-DEC-2000; 2000DE-1061338.
XX
XX (EPIG-) EPIGENOMICS AG.
XX
XX Schacht O;
PI
XX
XX WPI; 2002-500450/53.
DR
XX
XX New nucleic acid fragments from chemically treated
PT angiogenesis-associated genes, useful for determining methylation
PT status, e.g. in diagnosis or treatment of cancer -
XX
XX Claim 1; SEQ ID NO 89; 41pp + Sequence Listing; German.
XX
XX The invention relates to a nucleic acid (I) comprising a segment of 18
CC bases of chemically pretreated DNA of angiogenesis-associated genes (II)
CC having sequences (ABQ66971-ABQ67178) or their complements. (I), also
CC related oligomers, are used to evaluate the methylation status and/or
CC single-nucleotide polymorphisms, in angiogenesis-related genes, for
CC diagnosis and treatment of eye diseases, proliferative retinopathy,
CC neovascular glaucoma, solid tumours, inflammation, rheumatoid arthritis,
CC diabetic retinopathy, macular degeneration caused by neovascularisation,
CC psoriasis, arteriosclerosis, inflammatory bowel diseases, ulcers and
CC Crohn's disease.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 34688 BP; 9920 A; 444 C; 6941 G; 17383 T; 0 other;
SQ

XX DE Human single nucleotide polymorphism containing DNA sequence #1930.
XX KW Biallelic marker; polymorphism; human; disease; diagnosis; treatment;
XX KW phenotypic trait; gene therapy; forensic; paternity; mapping; cancer;
XX KW transgenic; single nucleotide polymorphism; SNP; ss.
XX OS Homo sapiens.
XX FT Key Location/Qualifiers
XX FT Variation replace(124,+)
XX FT /*tag= a
XX FT /standard_name= "single nucleotide polymorphism"
XX PN WO9953095-A2.
XX PD 21-OCT-1999
XX PF 30-MAR-1999; 99WO-US06893.
XX PR 09-APR-1998; 98US-0057871.
XX PA (WHED) WHITEHEAD INST BIOMEDICAL RES.
XX PI Lander ES, Wang D, Hudson T;
XX DR WPI; 1999-620443/53.
XX PT Polymorphic human genomic sequences and related allele-specific probes
XX PT and primers, useful for genetic analysis, e.g. diagnosis and monitoring
XX PT of disease
XX PS Claim 1; Page 235; 330pp; English.
XX CC This invention describes novel human nucleic acid segments (I)
XX CC containing polymorphic sites. The polynucleotides of (I) are used for,
XX CC e.g. correlating disease polymorphisms (or disease susceptibility) or
XX CC other phenotypic traits (e.g. baldness, obesity, fertility, strength,
XX CC response to drugs etc.); diagnosing and monitoring e.g. cancer,
XX CC inflammation, heart or central nervous system diseases; detecting
XX CC susceptibility to microbial infection; treating or preventing such
XX CC diseases; forensic analysis; gene therapy; paternity testing; mapping
XX CC genomic loci associated with phenotypic traits (and subsequent cloning
XX CC of the genes responsible); and the production of transgenic organisms.
XX CC Antibodies raised against (I) are useful as diagnostic and therapeutic
XX CC tools and in drug screening. AAH85144 - AAH87644 represent the human
XX CC DNA sequences containing biallelic polymorphic sites described in the
XX CC invention.
XX SQ Sequence 208 BP; 57 A; 58 C; 40 G; 53 T; 0 other;
Query Match 6.1%; Score 19; DB 20; Length 208;
Best Local Similarity 100.0%; Pred. No. 8-9;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 142 AATATACACCATATATAT 160
DB 43 AATATACACCATATATAT 61
RESULT 11
ID AAC07206/c
XX AAC07206 standard; cDNA; 307 BP.
XX AC AAC07206;
XX XX
XX 06-OCT-2000 (first entry)
XX DE Human secreted protein 5' EST, SEQ ID NO: 11281.
XX DE Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX KW gene therapy; chromosome mapping; ss.
XX XX

OS Homo sapiens.
XX EP1033401-A2.
XX PD 06-SEP-2000.
XX PF 21-FEB-2000; 2000EP-0200610.
XX PR 26-FEB-1999; 99US-0122487.
XX PA (GEST) GENSET.
XX PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX DR WPI; 2000-500381/45.
XX PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
XX PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX PS Claim 1; SEQ ID 11281; 71pp + CD-ROM; English.
XX CC The present sequence is one of a large number of 5' ESTs derived from
XX CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
XX CC identified within the present sequence. The 5' ESTs were prepared from
XX CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
XX CC sequences usually correspond mainly to the 3' untranslated region (UTR)
XX CC of the mRNA because they are often obtained from oligo-dT primed cDNA
XX CC libraries. Such ESTs are not well suited for isolating cDNA sequences
XX CC derived from the 5' ends of mRNAs and even in those cases where longer
XX CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
XX CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
XX CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
XX CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX CC They are used to obtain upstream regulatory sequences and to design
XX CC expression and secretion vectors.
XX SQ Sequence 307 BP; 91 A; 41 C; 68 G; 106 T; 1 other;
Query Match 6.1%; Score 19; DB 21; Length 307;
Best Local Similarity 100.0%; Pred. No. 8-9;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 225 AGAATTAGCTATTGTATC 243
DB 70 AGAATTAGCTATTGTATC 52
RESULT 12
AACX99087
ID AACX99087 standard; cDNA; 562 BP.
XX AC AACX99087;
XX XX
XX 24-SEP-1999 (first entry)
XX DE Human cancer cell derived cDNA contig #13.
XX KW Cancer; human; colon; breast; lung; transmembrane receptor; ATPase;
XX KW integral membrane protein; aspartyl protease; GATA family; wnt family;
XX KW transcription factor; G-protein alpha subunit; protein phosphatase;
XX KW phospholipase C; diacylglycerol binding protein; trypsin;
XX KW protein kinase; tyrosine phosphatase; developmental signalling protein;
XX KW WW/sp5/WWP domain; therapy; forensic; genetic mapping; diagnostic;
XX KW detection; treatment; cervical; melanoma; colorectal adenocarcinoma;
XX KW Wilms' tumour; retinoblastoma; sarcoma; myosarcoma; lung carcinoma;
XX KW leukemia; lymphoma; dysplasia; hyperplasia; endometrium; adrenal;
XX KW prostate; ss.
XX OS Homo sapiens.
XX XX
XX PN WO9933982-A2.
XX XX

PD 08-JUL-1999.

XX 22-DEC-1998; 98WO-US27610.

XX 21-DEC-1998; 98US-0217471.
 PR 23-DEC-1997; 97US-0068755.
 PR 03-APR-1998; 98US-0080664.
 PR 21-OCT-1998; 98US-0105234.
 PR 27-OCT-1998; 98US-0105877.

XX (CHIR) CHIRON CORP.

PA (HYSE-) HYSEQ INC.

PI Crkvenjakov R, Dickson M, Drmanac R, Drmanac S;

PI Escobedo J, Garcia PD, Garcia V, Giese K, Innis MA;

PI Jones LW, Kassam A, Kennedy GC, Kita D, Labat I;

PI Lamson G, Leshkowitz D, Pot D, Randazzo F, Reinhard C;

PI Stache-Crain B, Sudduth-Klinger J, Williams LT;

XX WPI; 1999-430243/36.

XX New isolated human polynucleotides

PS Claim 1; Page 578; 59lpp; English.

CC This invention describes novel isolated human polynucleotides obtained
 CC by screening for differential expression in colon cancer, breast cancer
 CC and lung cancer cell lines. The polynucleotides of the invention are
 CC represented in A98275-X9118 and encode polypeptides of protein
 CC families selected from 4 transmembrane segments integral membrane
 CC proteins, 7 transmembrane receptors, ATPases associated with various
 CC cellular activities (AAA), eukaryotic aspartyl proteases, GATA family of
 CC transcription factors, G-protein alpha subunit, phospholipase C,
 CC diacylglycerol binding proteins, protein kinase, phospholipase 2C,
 CC protein tyrosine phosphatase, trypsin, wnt family of developmental
 CC signalling proteins and WW/rsp5/WWP domain containing proteins. The
 CC encoded polypeptides also have a functional domain selected from Ank
 CC repeat, basic region plus leucine zipper transcription factors,
 CC bromodomain, EF-hand, SH3 domain, WD domain/G-beta repeats, zinc finger
 CC (C2H2 type), zinc finger (CCHC class), and zinc-binding metalloprotease
 CC domain. The polynucleotides encode polypeptides with similarity to known
 CC protein families and are predicted to have similar properties. The novel
 CC polynucleotides can be used to develop products for use as therapeutic
 CC agents and in forensics, genetic analysis, mapping and diagnostic
 CC applications. In particular, the product can be used for treating e.g. cervical
 CC and management of cancers. They can be used for treating e.g. cervical
 CC cancers, melanomas, colorectal adenocarcinomas, Wilms' tumour, sarcomas,
 CC retinoblastoma, myosarcomas, lung carcinomas, leukemias, such as chronic
 CC myelogenous leukemia, promyelocytic leukemia, monocytic leukemia, and
 CC myeloid leukemia, and lymphomas such as histiocytic lymphoma, anhydric
 CC hereditary ectodermal dysplasia, congenital alveolar dysplasia,
 CC epithelial dysplasia of the cervix, fibrous dysplasia of bone, and
 CC mammary dysplasia, hyperplasias, e.g. endometrial, adrenal, breast,
 CC prostate or thyroid hyperplasias or pseudoepitheliomatous hyperplasia of
 CC the skin.

XX Sequence 562 BP; 131 A; 152 C; 121 G; 146 T; 12 other;

Query Match 6.1%; Score 19; DB 20; Length 562;

Best-Local Similarity 100.0%; Pred.No. 9;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 142 AATATACACCATATATACT 160

DB 87 AATATACACCATATATACT 105

RESULT 13

AAH10915

ID AAH10915 standard; cDNA; 587 BP.

XX AAH10915;

AC AAH10915;

XX

DT 26-JUN-2001 (first entry)

XX Human cDNA clone (3'-primer) SEQ ID NO:7750.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

OS Homo sapiens.

PN EP1074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-0116126.

XX 29-JUL-1999; 99JP-0248036.

PR 27-AUG-1999; 99JP-0300253.

PR 11-JAN-2000; 2000JP-0118776.

PR 02-MAY-2000; 2000JP-0183767.

PR 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602
 CC full-length cDNAs defined in the specification, and for the detection
 CC and/or diagnosis of the abnormality of the proteins encoded by the
 CC full-length cDNAs -

PS Claim 3; SEQ ID 7750; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAH24446 to
 CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.

XX Sequence 587 BP; 157 A; 169 C; 104 G; 146 T; 11 other;

Query Match 6.1%; Score 19; DB 22; Length 587;

Best-Local Similarity 100.0%; Pred.No. 9;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 142 AATATACACCATATATACT 160

DB 76 AATATACACCATATATACT 94

RESULT 14

AZ98048/c

ID AAZ98048 standard; cDNA; 1164 BP.

XX

AC AAZ98048;

XX 09-MAY-2000 (first entry)
 XX Human secreted protein encoding nucleotide sequence SEQ ID NO:42.
 XX
 XX Human; secreted protein; diagnosis; cytostatic; immunosuppressive;
 XX antiinflammatory; nootropic; neuroprotective; antiallergic; cancer;
 XX tumour; neurodegenerative disorder; developmental abnormality; allergy;
 XX foetal deficiency; blood disorder; immune system disorder; arthritis;
 XX autoimmune disease; hepatic disease; renal disease; inflammation;
 XX Alzheimer's disease; behavioural disorder; schizophrenia; osteoporosis;
 XX infection; AIDS; spinal cord injury; transplant rejection; diabetes;
 XX asthma; sepsis; acne; psoriasis; cardiovascular disorder;
 XX reproductive disorder; gastrointestinal disorder; respiratory disorder;
 XX metabolic disorder; food additive; preservative; ss.
 XX
 XX Homo sapiens.
 XX
 XX WO200004140-A1.
 XX
 XX 27-JAN-2000. 99WO-US15849.
 XX
 XX 14-JUL-1999; 98US-0092921.
 XX
 XX 15-JUL-1998; 98US-0092922.
 XX
 XX 15-JUL-1998; 98US-0092956.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Ruben SM, Komatsoulis G, Duan RD, Rosen CA, Moore PA, Shi Y;
 XX Lafleur DW, Ebner R, Olsen HS, Brewer LA, Florence KA, Young PE;
 XX Mucenski M, Endress GA, Soppet DR;
 XX
 XX WPI: 2000-161128/14.
 XX P-PSDB; AAY87095.
 XX
 XX New isolated human genes, useful for diagnosis and treatment of, e.g.
 XX cancers, neurological or blood disorders -
 XX
 XX Claim 1; Page 328; 49app; English.
 XX
 XX The polynucleotide sequences given in AA298017 to AA298108 encode the
 XX human secreted proteins given in AAY87064 to AAY87223. Human secreted
 XX protein can have activities based on the tissues and cells the genes are
 XX expressed in. Examples of activities include: cytostatic;
 XX immunosuppressive; antiinflammatory; nootropic; neuroprotective; and
 XX antiallergic. The polynucleotides and their corresponding secreted
 XX polypeptides are useful for preventing, treating or ameliorating medical
 XX conditions, e.g. by protein or gene therapy. Also pathological conditions
 XX can be diagnosed by determining the amount of the new polypeptides in a
 XX sample or by determining the presence of mutations in the new
 XX polynucleotides. Human secreted protein s and their polynucleotides can
 XX be used for developing products for the diagnosis or treatment of cancer,
 XX tumours, neurodegenerative disorders, developmental abnormalities and
 XX foetal deficiencies, blood disorders, diseases of the immune system,
 XX autoimmune diseases, hepatic and renal disease, inflammation,
 XX allergies, Alzheimer's disease, behavioural disorders, schizophrenia,
 XX osteoporosis, arthritis, infections, AIDS, spinal cord injuries,
 XX transplant rejection, diabetes, asthma, sepsis, acne, psoriasis,
 XX cardiovascular disorders, reproductive disorders, gastrointestinal
 XX disorders, respiratory disorders and metabolic disorders. The
 XX proteins or polynucleotides can also be used as food additives or
 XX preservatives. The proteins are also useful for identifying their
 XX binding partners. AA298008 to AA298016 and AAY87063 are sequence used in
 XX the exemplification of the present invention.
 XX
 XX Sequence 1164 BP; 256 A; 322 C; 353 G; 231 T; 2 other;

Query Match 6.1%; Score 19; DB 21; Length 1164;
 Best Local Similarity 100.0%; Pred. No. 9;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 142 AATATACACCATATATACT 160
 DB 1065 AATATACACCATATATACT 1047
 RESULT 15
 AAD11661/C
 ID AAD11661 standard; cDNA; 1164 BP.
 AC AAD11661;
 XX 24-SEP-2001 (first entry)
 XX Human secreted protein-encoding gene 32 cDNA clone HBXGP60, SEQ ID NO:42.
 XX
 XX Human; secreted protein; proliferative disorder; cancer; tumour; asthma;
 XX foetal abnormality; developmental abnormality; haematopoietic disorder;
 XX immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
 XX Parkinson's disease; cognitive disorder; schizophrenia; skin disorder;
 XX psoriasis; sepsis; diabetes; atherosclerosis; cardiovascular disorder;
 XX inflammation; neurological disorder; Alzheimer's disease; food additive;
 XX angiogenic disorder; kidney disorder; gastrointestinal disorder; allergy;
 XX pregnancy-related disorder; endocrine disorder; infection; wound healing;
 XX cell culture; chemotaxis; vulnery; binding partner identification;
 XX gene therapy; ss.
 XX
 XX Homo sapiens.
 XX
 XX Key Location/Qualifiers
 XX CDS 143..310
 XX /tag= a
 XX /product= "Human secreted protein precursor"
 XX sig_peptide 143..208
 XX /tag= b
 XX mat_peptide 209..307
 XX /tag= c
 XX /product= "Mature human secreted protein"
 XX
 XX WO200151504-A1.
 XX
 XX 19-JUL-2001.
 XX
 XX 12-JAN-2001; 2001WO-US00911.
 XX
 XX 13-JAN-2000; 2000US-0482273.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Ruben SM, Komatsoulis GA, Duan DR, Rosen CA, Moore PA, Shi Y;
 XX Lafleur DW, Olsen HS, Brewer LA, Florence KA, Young PE, Soppet DR;
 XX Endress GA, Mucenski M, Ebner R;
 XX
 XX WPI: 2001-425865/45.
 XX P-PSDB; AAE06072.
 XX
 XX Isolated nucleic acid molecule encoding a human secreted protein is
 XX used in preventing, treating or ameliorating a medical condition -
 XX
 XX Claim 1; Page 695; 864pp; English.
 XX
 XX AAD11630-AAD11721 represent cDNAs corresponding to 71 human secreted
 XX protein genes, and AAE06041-AAE06132 represent the proteins they encode.
 XX AAE06133-AAE06205 represent human secreted protein fragments.
 XX The secreted proteins and their genes are useful for preventing, treating
 XX or ameliorating medical conditions, e.g., by protein or gene therapy.
 XX Pathological conditions can be diagnosed by determining the amount of the
 XX new protein in a sample or by determining the presence of mutations in
 XX the new genes. Specific uses are described for each of the 71 genes,
 XX based on the tissues in which they are most highly expressed, and include
 XX developing products for the diagnosis or treatment of proliferative
 XX disorders, cancer, tumours, foetal and developmental abnormalities,
 XX haematopoietic disorders, diseases of the immune system, AIDS, autoimmune
 XX diseases (e.g., rheumatoid arthritis), inflammation, allergies,

CC neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),
CC cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,
CC psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,
CC angiogenic disorders, kidney disorders, gastrointestinal disorders, the
CC pregnancy-related disorders, endocrine disorders, and infections. The
CC proteins can also be used to aid wound healing and epithelial cell
CC proliferation, to prevent skin aging due to sunburn, to maintain organs
CC before transplantation, for supporting cell culture of primary tissues,
CC to regenerate tissues, to identify their cognate ligands or binding
CC partners, and in chemotaxis, and can be used as a food additive or
CC preservative to modify storage properties. Antibodies specific for a
CC protein of the invention can be used in alleviating symptoms associated
CC with the disorders mentioned above, and in diagnostic immunoassays e.g.,
CC radioimmunoassay or enzyme linked immunosorbent assay (ELISA).
CC The present sequence represents a human secreted protein-encoding cDNA of
CC the invention.

XX
SQ Sequence 1164 BP; 256 A; 322 C; 353 G; 231 T; 2 other;

Query Match 6.1%; Score 19; DB 22; Length 1164;
Best Local Similarity 100.0%; Pred. No. 9;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 142 AATATACACCATATATACT 160
DB 1065 AATATACACCATATATACT 1047

Search completed: July 7, 2003, 19:22:16
Job time : 224 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run On: July 7, 2003, 18:15:30 ; Search time 1140 Seconds
(without alignments)
7939.450 Million cell updates/sec

Title: US-09-820-425B-12
Perfect score: 311
Sequence: 1 actccagctctgtgtgcaag.....gccattgcaggtaagctt 311

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.in.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.inv.*
- 32: em.htg.other.*
- 33: em.htg.mus.*
- 34: em.htg.pln.*
- 35: em.htg.rod.*
- 36: em.htg.mam.*
- 37: em.htg.vrt.*
- 38: em.sy.*
- 39: em.htgo.hum.*
- 40: em.htgo.mus.*
- 41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	311	100.0	311	6	AX259511	Sequence
2	305.2	98.5	179396	2	AC090107	Homo sapi
3	45.6	14.7	158548	3	PFMAL3P2	Al034558 Plasmodi
4	43.4	14.0	158156	2	AL357393	AL357393 Homo sapi
5	43.4	14.0	194079	9	AC020559	AC020559 Homo sapi
6	42.8	13.8	170973	9	AP000923	AP000923 Homo sapi
7	42.2	13.6	211133	2	AC124522	AC124522 Mus muscu
8	42.2	13.6	226278	2	AC126449	AC126449 Mus muscu
9	42	13.5	160451	9	AC123909	AC123909 Homo sapi
10	42	13.5	172012	9	AC009055	AC009055 Homo sapi
11	41.6	13.4	113810	9	AC068292	AC068292 Homo sapi
12	41.6	13.4	200208	9	AC013545	AC013545 Homo sapi
13	41.4	13.3	1831	9	AK027539	AK027539 Homo sapi
14	41.2	13.2	169816	2	AC118600	AC118600 Mus muscu
15	41.2	13.2	235733	2	AC105297	AC105297 Mus muscu
16	41	13.2	65971	9	AC013438	AC013438 Homo sapi
17	41	13.2	178238	2	AC016740	AC016740 Homo sapi
18	41	13.2	178238	2	AC016740	AC016740 Homo sapi
19	40.8	13.1	114070	2	AP005243	AP005243 Oryza sat
20	40.8	13.1	159191	2	AP005099	AP005099 Oryza sat
21	40.6	13.1	163415	9	AL359181	AL359181 Human DNA
22	40.4	13.0	6292	6	AX251494	AX251494 Sequence
23	40.4	13.0	152668	2	AC087623	AC087623 Homo sapi
24	40.2	12.9	95699	2	AC128749	AC128749 Rattus no
25	40.2	12.9	174615	9	AL445604	AL445604 Human DNA
26	40	12.9	63039	9	AC114780	AC114780 Homo sapi
27	40	12.9	76388	2	AC100200	AC100200 Mus muscu
28	40	12.9	87695	8	AP004579	AP004579 Lotus jap
29	40	12.9	110985	8	AP004578	AP004578 Lotus jap
30	40	12.9	149324	2	AL157822	AL157822 Homo sapi
31	40	12.9	234112	3	PFMALP2	AL035475 Plasmodi
32	39.8	12.8	110000	2	CY39B6_0	293599 Caenorhabdi
33	39.8	12.8	165942	9	AC104943	AC104943 Homo sapi
34	39.8	12.8	246237	3	CY39B6A	AL132948 Caenorhab
35	39.6	12.7	143379	2	AC130728	AC130728 Oryza sat
36	39.6	12.7	158286	2	AC111338	AC111338 Rattus no
37	39.6	12.7	176374	2	AC116056	AC116056 Rattus no
38	39.2	12.6	257817	2	AC006909	AC006909 Caenorhab
39	39	12.5	66552	9	AC124014	AC124014 Homo sapi
40	39	12.5	142199	2	AP004375	AP004375 Oryza sat
41	39	12.5	160401	9	AC009277	AC009277 Homo sapi
42	39	12.5	161638	2	AC096604	AC096604 Rattus no
43	39	12.5	178287	2	AC094211	AC094211 Rattus no
44	39	12.5	196406	2	AC107292	AC107292 Rattus no
45	38.8	12.5	94595	9	AL732591	AL732591 Human DNA

ALIGNMENTS

RESULT 1
AX259511
LOCUS AX259511 311 bp DNA linear PAT 26-OCT-2001
DEFINITION Sequence 12 from Patent WO0172775.
ACCESSION AX259511
VERSION AX259511.1 GI:16508668
KEYWORDS SOURCE
ORGANISM human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Chen,S.Y., Hu,P., Recipon,H. and Macina,R.A.
TITLE Compositions and methods of diagnosing, monitoring, staging,
imaging and treating lung cancer

JOURNAL Patent: WO 012775-A 12 04-OCT-2001;

diaDexus, Inc. (US)

FEATURES Location/Qualifiers

source 1..311

/organism="Homo sapiens"

/db_xref="taxon:9606"

BASE COUNT 107 a 54 c 64 g 86 t

ORIGIN

Query Match 100.0%; Score 311; DB 6; Length 311;

Best Local Similarity 100.0%; Pred. No. 7e-68;

Matches 311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTCCAGCTGTGTGCAAGGAGATGTCGGGAATGTCACAGCATGCTATAGCAAGAGC 60

|||||

Db 1 ACTCCAGCTGTGTGCAAGGAGATGTCGGGAATGTCACAGCATGCTATAGCAAGAGC 60

|||||

QY 61 ATATTGGCAACAGCTTGGATGGCCAGCAGAGAGGCCAAATGTGTGATTCATATTCAC 120

|||||

Db 61 ATATTGGCAACAGCTTGGATGGCCAGCAGAGAGGCCAAATGTGTGATTCATATTCAC 120

|||||

QY 121 AGTCGAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 180

|||||

Db 121 AGTCGAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 180

|||||

QY 181 TACTATAGTGTGATTCAGTTCACCTCCATTCAGTGGAAAAAATGGAAGAAATAGCTATT 240

|||||

Db 181 TACTATAGTGTGATTCAGTTCACCTCCATTCAGTGGAAAAAATGGAAGAAATAGCTATT 240

|||||

QY 241 ATCCATATGGGATACAAAAGCAGGGTACAAAAGATCTACATCTTGGCCATTGC 300

|||||

Db 241 ATCCATATGGGATACAAAAGCAGGGTACAAAAGATCTACATCTTGGCCATTGC 300

|||||

QY 301 AGCTAAAGCTT 311

|||||

Db 301 AGCTAAAGCTT 311

RESULT 2

AC090107/c

LOCUS

DEFINITION

Homo sapiens chromosome 12 clone RP11-643D8, WORKING DRAFT

SEQUENCE, 8 unordered pieces.

AC090107

AC090107.14 GI:20335798

HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.

KEYWORDS

Source

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 179396)

Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-Osman, F.R., Allen, C.,

Alsbrooks, S.L., Anaratinge, H.C., Are, J.R., Ayale, M., Banks, T.,

Barbieri, J., Benton, J., Bimaga, K., Blankenburg, K., Bonnin, D.,

Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P.,

Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,

Carton, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,

Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,

Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,

DeLaney, K.R., Davis, C., Davy-Carroll, L., Dederich, D.A.,

Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,

Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,

Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,

Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,

Correll, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,

Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,

Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B.,

Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,

Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,

Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,

Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,

Li, J., Li, Z., Lichtarge, O., Llieu, C., Liu, J., Liu, W., Louiseged, H.,

Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,
Massey, E., Mawhiney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M.,
Miner, G., Miner, Z., Mitchell, T., Monabbat, K., Morgan, M., Morris, S.,
Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
Nguyen, N., Nickerson, E., Nwokenkwo, S., Oguh, M., Okwuonu, G.,
Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
Rives, M., Rojas, A., Rojubokan, I., Rolfe, M., Ruiz, S., Savary, G.,
Scherer, S., Scott, G., Shen, H., Shooshari, N., Sisson, I.,
Sodergren, E., Sonalke, T., Sparks, A., Stanley, H., Stone, H.,
Sutton, A., Svatek, A., Taber, P., Tamerisa, A., Tamerisa, K., Tang, H.,
Tansley, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
Williams, G., Williamson, A., Wlarczyk, R., Wooden, S., Worley, K.,
Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G., and Gibbs, R.

Direct Submission

Unpublished

2 (bases 1 to 179396)

Worley, K.C.

Direct Submission

Submitted (15-FEB-2001)

Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 179396)

Worley, K.C.

Direct Submission

Submitted (09-MAY-2002)

Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

On Apr 28, 2002 this sequence version replaced gi:1844967.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: HCJT

Center clone name: RP11-643D8

----- Summary Statistics

Sequencing vector: Plasmid;

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 175679 bases at least Q40

Consensus quality: 176511 bases at least Q30

Consensus quality: 176995 bases at least Q20

Estimated insert size: 176936; sum-of-contigs estimation

Quality coverage: 10x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length

(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

consists of 8 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

as soon as it is available and the accession number will

be preserved.

* 1 4459: contig of 4459 bp in length

* 4460 4559: gap of unknown length

* 4560 9953: contig of 5394 bp in length

* 9954 10053: gap of unknown length

* 10054 17744: contig of 7691 bp in length

* 17745 17844: gap of unknown length

* 17845 25913: contig of 8069 bp in length

* 25914 26013: gap of unknown length

* 26014 41673: contig of 15660 bp in length

* 41674 41773: gap of unknown length

* 41774 70908: contig of 29135 bp in length

* 70909 71008: gap of unknown length

```
* 71009 121798: contig of 50790 bp in length
* 121799 121898: gap of unknown length
* 121899 179396: contig of 57498 bp in length.

FEATURES             Location/Qualifiers
     source           1..179396
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="12"
                     /clone="RP11-643D8"
BASE COUNT  50740 a 38957 c 38965 g 50021 t      713 others
ORIGIN

Query Match      98.5%; Score 306.2; DB 2; Length 179396;
Best Local Similarity 99.0%; Pred. No. 9.3e-67;
Matches 308; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ACTCCAGCTCTGTGTGCAAGAGATGCTGCTGGAATGTCACAGATCGTATAGCAAGAGC 60
    |||||||
Db 138531 ACTCCAGCTCTGTGTGCAAGAGATGCTGCTGGAATGTCACAGATCGTATAGCAAGAGC 138472
    |||||||

QY 61 ATATTGGCAACAGCTTGGATGCCAGCAGAGAGGCCAACAAATGCTGATTCATATTCAC 120
    |||||||
Db 138471 ATATTGGCAACAGCTTGGATGCCAGCAGAGAGGCCAACAAATGCTGATTCATATTCAC 138412
    |||||||

QY 121 AGTCGAATPAATGAATACATAATACATATACACATATATACACTGTATGTTGTTCTA 180
    |||||||
Db 138411 AGTCGAATPAATGAATACATAATACATATATACACTGTATGTTGTTCTA 138352
    |||||||

QY 181 TACTATAGTATTGACTTGAATCACTCAATCACTGAAAGAAATGGAAGATAGCTATTTGT 240
    |||||||
Db 138351 TACTATAGTATTGACTTGAATCACTCAATCACTGAAAGAAATGGAAGATAGCTATTTGT 138292
    |||||||

QY 241 ATCCATATGGGTACAAAAAGAGCGGTAACAAAGAAATCTACATCATCTTGCCTATTTGC 300
    |||||||
Db 138291 ATCCATATGGGTACAAAAAGAGCGGTAACAAAGAAATCTACATCATCTTGCCTATTTGC 138232
    |||||||

QY 301 AGGTAAGCTT 311
    |||||||
Db 138231 AGGTAAGCTT 138221
    |||||||

RESULT 3
PFMAL3P2/c          158548 bp      DNA      linear      INV 26-APR-2000
LOCUS               Plasmodium falciparum MAL3P2, complete sequence.
DEFINITION          AL034558 AL008982 AL008989 AL009007 AL009008 AL009009 AL009014
ACCESSION           AL010140 AL010150 AL010152 AL010156 AL010158 AL010211 AL022217
                    297349 297350 298546
VERSION             AL034558.3 GI:15375364
KEYWORDS            HTG.
SOURCE              Plasmodium falciparum 3D7.
ORGANISM            Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE           1 (bases 1 to 158548)
AUTHORS             Bowman, S., Lawson, D., Basham, D., Brown, D., Chillingworth, T.,
                    Churcher, C. M., Craig, A., Davies, R. M., Devlin, K., Feltwell, T.,
                    Gentles, S., Gilliam, R., Hamlin, N., Harris, D., Holroyd, S.,
                    Hornsby, T., Horrocks, P., Jagels, K., Jassal, B., Kyes, S., McLean, J.,
                    Moule, S., Mungall, K., Murphy, L., Oliver, K., Quail, M. A.,
                    Rajandream, M. A., Rutter, S., Skelton, J., Squares, R., Squares, S.,
                    Sulston, J. E., Whitehead, S., Woodward, J. R., Newbold, C. and
                    Barrell, B. G.
TITLE               The complete nucleotide sequence of chromosome 3 of Plasmodium
                    falciparum
JOURNAL             Nature 400 (6744), 532-538 (1999)
MEDLINE            99376085
PUBMED            10448855
REFERENCE           2 (bases 1 to 158548)
AUTHORS             Churcher, C., Bowman, S., Lawson, D., Quail, M. and Barrell, B.
JOURNAL            Unpublished
REFERENCE           3 (bases 1 to 158548)
AUTHORS             Lawson, D., Bowman, S. and Barrell, B.
TITLE              Direct Submission

JOURNAL Submitted (17-DEC-1998) P.falciparum Genome Sequencing Consortium,
The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge
CB10 1SA, UK
COMMENT On or before Aug 30, 2001 this sequence version replaced
gi:2982545, gi:2982529, gi:2982547, gi:2982546, gi:2982548,
gi:2624372, gi:2982553, gi:2982560, gi:2894487, gi:2665315,
gi:2982565, gi:2982575, gi:2982578, gi:2982517, gi:2982518,
gi:2982519, gi:4493878.
For more information about this sequence or the Malaria Project,
see http://www.sanger.ac.uk/Projects/P_falciparum.
FEATURES             Location/Qualifiers
     source           1..158548
                     /organism="Plasmodium falciparum 3D7"
                     /strain="3D7"
                     /db_xref="taxon:36329"
                     /chromosome="3"
                     /clone="MAL3P2"
                     885..898
                     /note="cacaccccccccccc sequence upstream of PFC0160w"
                     1011..1018
                     /note="C8-tract upstream of PFC0160w"
                     join(1030..2337,2429..2502,2645..2685,2826..2941,
                     3046..3173,3308..3388,3508..3604,3699..3737)
                     /gene="PFC0160w, MAL3P2.1"
                     join(1030..2337,2429..2502,2645..2685,2826..2941,
                     3046..3173,3308..3388,3508..3604,3699..3737)
                     /gene="PFC0160w, MAL3P2.1"
                     /note="PFC0160w, (MAL3P2.1), Putative Ankyrin repeat
                     protein, len: 627 aa, Pfam match (x4) to entry PF00023
                     ank, Ank repeat, Pfam match to entry PF01529 zf-DHHC, DHHC
                     zinc finger domain, revised: revised to 8 exon structure"
                     /codon_start=1
                     /product="putative Ankyrin repeat protein"
                     /protein_id="CAB38988.2"
                     /db_xref="GI:15375365"
                     /db_xref="SPTREMBL:O97224"
                     /translation="MNDNESLSEVLEKOEYIIRKQYDFIRQLIOPYLLNNDIE
                     MNSINILHWACYCGFTLKVRLKIFNCDIEKIDVNDNTPIYAIRKSNYEIVILLI
                     KCFGISILFHKRRRMSPTLAICEFNEDKILEALHLELLVMGVSLEONEHGOTA
                     LFLSVKNNISTLOWLLTKEVINHRDFYGNVLHIAVRHCDIDILRLICDYGCLNMV
                     YSSIEKNNTWVQLCIKNRFLYILLKKVQLNKCICKICKIYAFYFWFFAIL
                     NLIVFNIAHSFSLINKYHFKSLIWTITWFFQFLWCMLEFKSPGFYKNNHMFNNKK
                     NIQSKYNSNPMYNGTPTNAEYQLNNIEREIQINKLLISTNFTPTPINHDQALS
                     NKYNDILINLEYQLSLYSQVQERINSLDEIRNAILNQNPRICVTCNLIKPPRV
                     HHCACGHCIVHQDHHCWVDNCIGIKNRCFYMFICIFVLLLYNYVYLYFHLFQ
                     ATINAFGSLVILGNCFINVTFAITLYLFARNTILTNVTFYEHVKPNHTDKYNT
                     DIRGWDQNLNLKGLIKNIYFWSLNDPEYIRQCTKTVINLDTYKNSNDVCYTLI
                     SNT"
                     2338..2343
                     /gene="PFC0160w, MAL3P2.1"
                     /note="potential donor sequence for exon 1 of PFC0160w
                     (revised)"
                     2416..2428
                     /gene="PFC0160w, MAL3P2.1"
                     /note="potential acceptor sequence for exon 2 PFC0160w
                     (revised)"
                     join(2435..2502,2645..2685,2826..2878)
                     /gene="PFC0160w, MAL3P2.1"
                     /note="Match to PF01529 zf-DHHC, DHHC zinc finger domain
                     Score 58.99"
                     2503..2508
                     /gene="PFC0160w, MAL3P2.1"
                     /note="potential donor sequence for exon 2 of PFC0160w
                     (revised)"
                     2633..2644
                     /gene="PFC0160w, MAL3P2.1"
                     /note="potential acceptor sequence for exon 3 of PFC0160w
                     (revised)"
                     2686..2691
                     /gene="PFC0160w, MAL3P2.1"
                     /note="potential donor sequence for exon 3 of PFC0160w
                     (revised)"
                     2816..2825
```


[illegible]

```
* 58762 58861: gap of unknown length
* 58862 120891: contig of 62030 bp in length
* 120892 120991: gap of unknown length
* 120992 211133: contig of 90142 bp in length.
FEATURES
  source      Location/Qualifiers
    1..211133
      /organism="Mus musculus"
      /db_xref="taxon:10090"
      /chromosome="UNK"
      /clone="RP23-397A19"
    1..1884
      /note="assembly_name:Contig14"
      /note="assembly_name:Contig15"
      /note="assembly_name:Contig16"
      /note="assembly_name:Contig17"
      /note="assembly_name:Contig18"
      /note="assembly_name:Contig19"
BASE COUNT  63603 a 42064 c 42250 g 62816 t 400 others
ORIGIN
Query Match      13.6%; Score 42.2; DB 2; Length 211133;
Best Local Similarity 53.1%; Pred. No. 1.6;
Matches 112; Conservative 0; Mismatches 98; Indels 1; Gaps 1;
QY 96 CCCAAATGCTGATTCATATTCCTAGTTCGAATTAATGATACATACATATACACCATAT 155
Db 70155 CCAACTGTTAAATGAAGCAATATATATATATATATATATATATATATATATAT 70096
QY 156 ATACTAGACTGTATGTGTTCTTACTATATAGTGTGACTTGAACCTCAATTCAGTGAA 215
Db 70095 ATACTATATATGTTGTGTATATATATATATATATATATATATATATATATATAT 70036
QY 216 AAAATGGAAGA-ATTAGCTATTTTATCCATATGGGATACAAAAAGAGGTAACAA 274
Db 70035 AATTTGAAAATATTCTACTGATTTGACCTGTTTCAGAGTCACACATTTAGAGGAA 69976
QY 275 AGAATCTACATCATCTTGGCATTTGCAGGTA 305
Db 69975 AGAATTAATCTAAATCTCAGTGGTAGATA 69945
RESULT 8
AC126449
LOCUS      226278 bp DNA linear HTG 05-JUL-2002
DEFINITION Mus musculus chromosome UNK clone RP23-359P11, WORKING DRAFT
ACCESSION  AC126449
VERSION     GI:21699739
KEYWORDS   HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE     house mouse.
ORGANISM   Mus musculus
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 226278)
AUTHORS   McPherson,J.D. and Waterston,R.H.
TITLE     The sequence of Mus musculus clone
JOURNAL   Unpublished
REFERENCE  2 (bases 1 to 226278)
AUTHORS   McPherson,J.D. and Waterston,R.H.
TITLE     Direct Submission
JOURNAL   Submitted (05-JUL-2002) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
COMMENT
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@watson.wustl.edu
----- Project Information -----
Center project name: M_BA0359P11
```

```
----- Summary Statistics -----
Sequencing vector: M13; 0%
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 222728 bases at least Q40
Consensus quality: 223933 bases at least Q30
Consensus quality: 224684 bases at least Q20
* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1219: contig of 1219 bp in length
* 1220 1319: gap of unknown length
* 1320 2577: contig of 1258 bp in length
* 2578 2677: gap of unknown length
* 2678 28502: contig of 25825 bp in length
* 28503 28602: gap of unknown length
* 28603 62839: contig of 34237 bp in length
* 62840 62939: gap of unknown length
* 62940 226278: contig of 163339 bp in length.
FEATURES
  source      Location/Qualifiers
    1..226278
      /organism="Mus musculus"
      /db_xref="taxon:10090"
      /chromosome="UNK"
      /clone="RP23-359P11"
    1..1219
      /note="assembly_name:Contig3"
      /note="assembly_name:Contig4"
      /note="assembly_name:Contig6"
      /note="assembly_name:Contig7"
      /note="assembly_name:Contig8"
      /note="assembly_name:Contig9"
BASE COUNT  70700 a 42334 c 42443 g 70401 t 400 others
ORIGIN
Query Match      13.6%; Score 42.2; DB 2; Length 226278;
Best Local Similarity 49.8%; Pred. No. 1.6;
Matches 107; Conservative 0; Mismatches 108; Indels 0; Gaps 0;
QY 93 GAGCCCAATGCTGATTCATATTCCTAGTTCGAATTAATGATACATACATACACCA 152
Db 221420 GAGGATATTGAATATTATTTCTTTATTTATATATTCAGTTAGTATAGTATACACTA 221479
QY 153 TATATACTAGACTGTATGTGTTCTTCTATCTATCTATCTATCTATCTATCTATCTAT 212
Db 221480 TGTTTAATTTCTGATCTGCTGTTAGCTTCGTCGAATGATTTCTTATATGCAATGCA 221539
QY 213 GAAAAAATGGAAGAATTAGCTATTTTGTATCCATATGGGATACAAAAAGAGGGTAACA 272
Db 221540 TGCTGCAAGTTTATTTTGGTATTATTTTATATATATCGTAGAAAAAATCCATCTCTG 221599
QY 273 AAAGAATCTACATCATCTTCCTATTCAGGTTAA 307
Db 221600 GTGTACTACATTTCTCTATCCATTTTCAGTTAA 221634
RESULT 9
AC123909
LOCUS      160451 bp DNA linear PRI 05-JUN-2002
DEFINITION Homo sapiens chromosome 16 clone RP11-27L11, complete sequence.
ACCESSION  AC123909 AC025481
```



```

QY      228 ATTAGCTATTGTATCATGGATACAAAGGAGCGGTAACAAAGAAATCATCATCA 287
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      168044 AATACACAATATATTGTATTGTATTAACATATATACAAACACAGTATTGTGTTG 167985
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      288 TCTTGCCATTGCA 301
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      167984 TATTACTATATACA 167971
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

RESULT 11
AC068292/c
LOCUS      AC068292      113810 bp      DNA      linear      PRI 07-NOV-2001
DEFINITION Homo sapiens BAC clone RP11-738L3 from 2, complete sequence.
ACCESSION AC068292
VERSION   AC068292.8 GI:13625515
KEYWORDS  HTG.
SOURCE    Homo sapiens.
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 113810)
           Sulston, J.E. and Waterston, R.
           Toward a complete human genome sequence
           Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE   99063792
PUBMED    9847074
REFERENCE 2 (bases 1 to 113810)
           Isak, A. and Haakenson, W.
           The sequence of Homo sapiens BAC clone RP11-738L3
           Unpublished
REFERENCE 3 (bases 1 to 113810)
           Waterston, R.H.
           Direct Submission
           Submitted (30-APR-2000) Genome Sequencing Center, Washington
           University School of Medicine, 4444 Forest Park Parkway, St. Louis,
           MO 63108, USA
REFERENCE 4 (bases 1 to 113810)
           Waterston, R.H.
           Direct Submission
           Submitted (15-APR-2001) Genome Sequencing Center, Washington
           University School of Medicine, 4444 Forest Park Parkway, St. Louis,
           MO 63108, USA
REFERENCE 5 (bases 1 to 113810)
           Waterston, R.
           Direct Submission
           Submitted (09-AUG-2001) Department of Genetics, Washington
           University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
           6 (bases 1 to 113810)
           Waterston, R.
           Direct Submission
           Submitted (07-NOV-2001) Department of Genetics, Washington
           University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
           On Apr 15, 2001 this sequence version replaced gi:13518289.
           ----- Genome Center
           Center: Washington University Genome Sequencing Center
           Center code: WUGSC
           Web site: http://genome.wustl.edu/gsc
           Contact: sapiens@watson.wustl.edu
           ----- Summary Statistics
           -----
           Center project name: H_NH0738L03
           -----

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by

restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frengen, E., Tatenno, M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (vector: pBACpac.med.buffalo.edu) (<http://bacpac.med.buffalo.edu>)

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-86A21, 200 bp overlap; the clone sequenced to the right is RP11-337M10. Actual start of this clone is at base position 145439 of RP11-86A21; actual end is at base position 113810 of RP11-738L3.

FEATURES	Location/Qualifiers	Source
repeat_region	12..1117	/organism="Homo sapiens"
repeat_region	1119..1343	/db_xref="taxon:9606"
repeat_region	1627..1705	/chromosome="2"
repeat_region	2315..3250	/map="2"
repeat_region	3284..3372	/clone="RP11-738L3"
repeat_region	3373..3724	/clone_lib="RPCI-11"
repeat_region	3725..4027	/rpt_family="L1"
repeat_region	4035..5056	/rpt_family="A-rich"
repeat_region	5057..5357	/rpt_family="L1"
repeat_region	5358..6500	/rpt_family="ERVL"
repeat_region	7110..7281	/rpt_family="MALR"
repeat_region	7501..7678	/rpt_family="ERVL"
repeat_region	8353..8497	/rpt_family="L1"
repeat_region	8665..8969	/rpt_family="Alu"
repeat_region	10285..10398	/rpt_family="L1"
repeat_region	10442..10514	/rpt_family="L2"
repeat_region	10991..11011	/rpt_family="MIR"
repeat_region	11162..11280	/rpt_family="AT-rich"
repeat_region	13231..13326	/rpt_family="MIR"
repeat_region	13945..14320	/rpt_family="L2"
repeat_region	14315..14387	/rpt_family="L1"
repeat_region		/rpt_family="L1"

OK 73019, USA
On Aug 10, 2002 this sequence version replaced gl:21956588.
COMMENT
----- Genome Center

COMMENT

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 7, 2003, 19:18:29 ; Search time 135 Seconds
(without alignments)
3596.062 Million cell updates/sec

Title: US-09-820-425B-12
Perfect score: 311
Sequence: 1 actccagctctgtgcaag.....gccattgcaggtaaagctt 311

Scoring table: OLIGO.NUC

Gapop 60.0 , Gapext 60.0

Searched: 1085931 seqs, 780495707 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2171862

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published_Applications_NA.*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	311	100.0	311	10	US-09-820-425B-12
2	29	9.3	29	10	US-09-820-425B-19
3	22	7.1	22	10	US-09-820-425B-17
4	19	6.1	601	10	US-09-777-921A-48
5	19	6.1	601	10	US-09-777-921A-49
6	19	6.1	1164	9	US-09-984-271-42
7	19	6.1	2273	12	US-10-062-254-101
8	19	6.1	69327	10	US-09-777-921A-3
9	18	5.8	18	10	US-09-820-425B-18
10	18	5.8	1465	10	US-09-864-761-13828
11	18	5.8	1521	10	US-09-864-761-30395
12	18	5.8	3697	9	US-10-205-823-368
13	18	5.8	4101	9	US-09-981-353-28
14	18	5.8	32192	9	US-10-092-154-1416
15	18	5.8	32192	9	US-09-764-891-7945
16	18	5.8	32192	9	US-09-764-891-8220
17	18	5.8	32192	9	US-09-764-891-8319
18	18	5.8	32192	10	US-09-764-847-1416
19	17	5.5	28	10	US-09-774-107-12

ALIGNMENTS

RESULT 1
US-09-820-425B-12
; Sequence 12, Application US/09820425B
; Patent No. US20020119156A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Sel-Yu
; APPLICANT: Hu, Ping
; APPLICANT: Recipon, Herve
; APPLICANT: Macina, Roberto
; TITLE OF INVENTION: Compositions and Methods of Diagnosing, Monitoring,
; FILE REFERENCE: Staging, Imaging and Treating Lung Cancer
; FILE REFERENCE: DEX-0203
; CURRENT APPLICATION NUMBER: US/09/820,425B
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/192,921
; PRIOR FILING DATE: 2000-03-29
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 311
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-820-425B-12

Query Match	100.0%	Score 311;	DB 10;	Length 311;
Best Local Similarity	100.0%	Pred. No. 1.1e-155;		
Matches 311;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY 1 ACTCCAGCTCTGTGTGCAAGGAGATGTGCTGGAATGTACACATCGTATAGCAAGAGC 60
|||||
Db 1 ACTCCAGCTCTGTGTGCAAGGAGATGTGCTGGAATGTACACATCGTATAGCAAGAGC 60
|||||
QY 61 ATATTGGCAACAGCTTGGATGCCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
|||||
Db 61 ATATTGGCAACAGCTTGGATGCCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
|||||
QY 121 ACTCGAATAATTGAATACTACCAATATACACATATATAGACTGTATGTTGTTCTA 180
|||||
Db 121 ACTCGAATAATTGAATACTACCAATATATAGACTGTATGTTGTTCTA 180
|||||
QY 181 TACTATAGTGTGTTGACTTGAACCTCCATTGAGTAAAGAAATAGCTATTTGT 240
|||||

	Db	181	TACTATAGTGATTGACTTTGAACCTCCATTCACTGAGAAAAAATGGAGAATAATGCTATTTCGT	240
	Qy	241	ATCCCATATGGGATACAAAAGCAGGGTACAAAGAAATCTACATCATCTGTGCCATTTCG	300
	Db	241	ATCCCATATGGGATACAAAAGCAGGGTACAAAGAAATCTACATCATCTGTGCCATTTCG	300
	Qy	301	AGGTAAAGCTT	311
	Db	301	AGGTAAAGCTT	311

RESULT 2

```

US-09-820-425B-19
; Sequence 19, Application US/09820425B
; Patent No. US20020119156A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Sei-Yu
; APPLICANT: Hu, ping
; APPLICANT: Recipon, Herve
; APPLICANT: Macina, Roberto
; TITLE OF INVENTION: Compositions and Methods of Diagnosing, Monitoring,
; FILE REFERENCE: Staging, Imaging and Treating Lung Cancer
; FILE REFERENCE: DEX-0203
; CURRENT APPLICATION NUMBER: US/09/820,425B
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/192,921
; PRIOR FILING DATE: 2000-03-29
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 29
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-820-425B-19

```

```
Query Match      9.3%; Score 29; DB 10; Length 29;
Best Local Similarity 100.0%; Pred. No. 1.9e-05;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

RESULT 3

```

US-09-820-425B-17
; Sequence 17, Application US/09820425B
; Patent No. US20020119156A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Sei-Yu
; APPLICANT: Hu, Ping
; APPLICANT: Recipon, Hervé
; APPLICANT: Macina, Roberto
; TITLE OF INVENTION: Compositions and Methods of Diagnosing, Monitoring,
; TITLE OF INVENTION: Staging, Imaging and Treating Lung Cancer
; FILE REFERENCE: DEX-0203
; CURRENT APPLICATION NUMBER: US/09/820,425B
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/192,921
; PRIOR FILING DATE: 2000-03-29
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-820-425B-17

```

Query Match	7.18; Score 22; DB 10; Length 22;
-------------	-----------------------------------

```
Best Local Similarity 100.0%; Pred. No. 0.1;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 17 CAAGGAGATGTCTGGAAATGTC 38
|||||||
Db 1 CAAGGAGATGTCTGGAAATGTC 22

RESULT 4

```

US-09-777-921A-48
; Sequence 48, Application US/09777921A
; Patent No. US20020115136A1
; GENERAL INFORMATION:
; APPLICANT: MERKULOV et al.
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL001103
; CURRENT APPLICATION NUMBER: US/09/777,921A
; CURRENT FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 48
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-777-921A-48

```

Query Match	6.18;	Score 19;	DB 10;	Length 601;
Best Local Similarity	100.0%;	Pred. No.	4.9;	
Matches 19;	Conservative	0;	Mismatches	0; Indels
Gaps	0;			
Qy	212	TGAAAAAATCGAGAATT	230	
Db	577	TGAAAAAATCGAGAATT	595	

RESULT 5

```

US-09-777-921A-49
; Sequence 49, Application US/09777921A
; Patent No. US20020115136A1
; GENERAL INFORMATION:
; APPLICANT: MERKULOV et al.
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL001103
; CURRENT APPLICATION NUMBER: US/09/777, 921A
; CURRENT FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 49
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-777-921A-49

```

Query Match	6.1%	Score 19;	DB 10;	Length 601;
Best Local Similarity	100.0%;	Pred. No. 4.9;		
Matches 19;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;

RESULT 6

US-09-984-271-42/c
; Sequence 42, Application US/09984271
; Publication No. US2003004008A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 71 Human Secreted Proteins

FILE REFERENCE: PZ030P1
; CURRENT APPLICATION NUMBER: US/09/984,271
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 09/482,273
; PRIOR FILING DATE: 2000-01-13
; PRIOR APPLICATION NUMBER: PCT/US99/15849
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: 60/092,921
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/092,922
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/092,956
; PRIOR FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 267
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 42
; LENGTH: 1164
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (582)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (592)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-984-271-42

Query Match 6.1%; Score 19; DB 9; Length 1164;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 142 AATATACACCATATATACT 160
|||||
DB 1065 AATATACACCATATATACT 1047

RESULT 7

US-10-062-254-101/c
; Sequence 101, Application US/10062254
; Patent No. US2002013882A1
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Edgar B
; APPLICANT: Cahoon, Rebecca E
; APPLICANT: Falco, Savio Carl
; APPLICANT: Fang, Yiwen
; APPLICANT: Hantke, Sabine S.
; APPLICANT: Lee, Jian-Ming
; APPLICANT: Li, Zhongsen
; APPLICANT: Miao, Guo-Hua
; APPLICANT: Morgante, Michele
; APPLICANT: Niu, Xiping
; APPLICANT: Odell, Joan
; APPLICANT: Rafalski, Antoni
; APPLICANT: Sakai, Hajime
; APPLICANT: Zheng, Peizhong
; APPLICANT: Zhu, Qun
; TITLE OF INVENTION: Polynucleotides Encoding Proteins Involved In Plant Metabolism
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/062,254
; CURRENT FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 09/630,346
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 60/146511
; PRIOR FILING DATE: 1999-07-30
; PRIOR APPLICATION NUMBER: 60/156006
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/156899
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: 60/157287
; PRIOR FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: 60/169767
; PRIOR FILING DATE: 1999-12-09

; PRIOR APPLICATION NUMBER: 60/171054
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: 60/172958
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 60/171515
; PRIOR FILING DATE: 1999-12-22
; PRIOR APPLICATION NUMBER: 60/173535
; PRIOR FILING DATE: 1999-12-29
; NUMBER OF SEQ ID NOS: 375
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 101
; LENGTH: 2273
; TYPE: DNA
; ORGANISM: Glycine max
; US-10-062-254-101

Query Match 6.1%; Score 19; DB 12; Length 2273;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 283 CATCATCTGCCATTGCA 301
|||||
DB 1283 CATCATCTGCCATTGCA 1265

RESULT 8

US-09-777-921A-3
; Sequence 3, Application US/09777921A
; Patent No. US20020115136A1
; GENERAL INFORMATION:
; APPLICANT: MERKULOV et al.
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL001103
; CURRENT APPLICATION NUMBER: US/09/777,921A
; CURRENT FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 69327
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(69327)
; OTHER INFORMATION: n = A,T,C or G
US-09-777-921A-3

Query Match 6.1%; Score 19; DB 10; Length 69327;
Best Local Similarity 100.0%; Pred. No. 6.5;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 212 TCAAAAAAATGGAAGAAATT 230
|||||
DB 6809 TCAAAAAAATGGAAGAAATT 6827

RESULT 9

US-09-820-425B-18/c
; Sequence 18, Application US/09820425B
; Patent No. US20020119156A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Sei-Yu
; APPLICANT: Hu, Ping
; APPLICANT: Recipon, Herve
; APPLICANT: Macina, Roberto
; TITLE OF INVENTION: Compositions and Methods of Diagnosing, Monitoring,
; TITLE OF INVENTION: Staging, Imaging and Treating Lung Cancer
; FILE REFERENCE: DEX-0203
; CURRENT APPLICATION NUMBER: US/09/820,425B
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/1192,921

```
; PRIOR FILING DATE: 2000-03-29
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 18
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-820-425B-18

Query Match          5.8%; Score 18; DB 10; Length 18;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 CCACGACGAGGCCCAA 100
    |||||||
Db 18 CCACGACGAGGCCCAA 1

RESULT 10
US-09-864-761-13828
; Sequence 13828, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecomica-x-1
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 13828
```

```
; LENGTH: 1465
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC011406.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 5.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.97
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.93
; OTHER INFORMATION: EXPRESSED IN HELI00, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.5
US-09-864-761-13828

Query Match          5.8%; Score 18; DB 10; Length 1465;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 ATACTATAGTGATTGACT 197
    |||||||
Db 1358 ATACTATAGTGATTGACT 1375

RESULT 11
US-09-864-761-30395
; Sequence 30395, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aecomica-x-1
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
```

NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 30395
; LENGTH: 1521
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC011406.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 5.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.97
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.93
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.5
; OTHER INFORMATION: NT HIT: g114146769, EVALUE 0.00e+00
; OTHER INFORMATION: EST HUMAN HIT: AI557282.1, EVALUE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: P50443, EVALUE 0.00e+00
US-09-864-761-30395

Query Match 5.8%; Score 18; DB 10; Length 1521;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 180 ATACTATAGTGATTGACT 197
|||||
Db 1241 ATACTATAGTGATTGACT 1258

RESULT 12

US-10-205-823-368
; Sequence 368, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Gorbacheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Womsey, Angela M.
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 368
; LENGTH: 3697
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-205-823-368

Query Match 5.8%; Score 18; DB 9; Length 3697;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 180 ATACTATAGTGATTGACT 197
|||||
Db 2188 ATACTATAGTGATTGACT 2205

RESULT 13

US-09-981-353-28
; Sequence 28, Application US/09981353
; Patent No. US20020160382A1
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy W.
; APPLICANT: Jones, David A.
; TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
; FILE REFERENCE: PA-0038 US
; CURRENT APPLICATION NUMBER: US/09/981.353
; CURRENT FILING DATE: 2001-10-11
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 28
; LENGTH: 4101
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; OTHER INFORMATION: Incyte ID No. US20020160382A1 333076.1
US-09-981-353-28

Query Match 5.8%; Score 18; DB 9; Length 4101;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 180 ATACTATAGTGATTGACT 197
|||||
Db 2163 ATACTATAGTGATTGACT 2180

RESULT 14

US-10-092-154-1416
; Sequence 1416, Application US/10092154
; Publication No. US20030054375A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009C1
; CURRENT APPLICATION NUMBER: US/10/092,154
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 2003
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1416
; LENGTH: 32192
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-092-154-1416

Query Match 5.8%; Score 18; DB 9; Length 32192;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 GTGTGCAAGGAGATGTC 29
|||||
Db 4920 GTGTGCAAGGAGATGTC 4937

RESULT 15

US-09-764-891-7945/c
; Sequence 7945, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891

; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7945
; LENGTH: 32192
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-7945

Query Match 5.88; Score 18; DB 9; Length 32192;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 GTGTGCAAGGAGATGTGC 29
|||||
DB 27273 GTGTGCAAGGAGATGTGC 27256

Search completed: July 7, 2003, 20:09:16
Job time : 137 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 7, 2003, 18:50:44 ; Search time 1134 Seconds
(without alignments)
7981.457 Million cell updates/sec

Title: US-09-820-425b-12

Perfect score: 311

Sequence: 1 actccagctctgtgtgcaag.....gccatttcgaggtaaagctt 311

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 2054640 seqs, 14551402878 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*
2: gb_hgt.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vi.*
30: em_hgt_hum.*
31: em_hgt_inv.*
32: em_hgt_other.*
33: em_hgt_mus.*
34: em_hgt_pln.*
35: em_hgt_rtd.*
36: em_hgt_mam.*
37: em_hgt_vrt.*
38: em_sy.*
39: em_hgtgo_hum.*
40: em_hgtgo_mus.*
41: em_hgtgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	311	100.0	311	6	AX259511	Sequence
2	158	50.8	179396	2	AC090107	Homo sapi
3	29	9.3	29	6	AX259518	Sequence
4	22	7.1	22	6	AX259516	Sequence
5	21	6.8	4373	9	AF214528S1	Homo sapi
6	21	6.8	7110	6	AX251243	Sequence
7	21	6.8	7110	6	AX251990	Sequence
8	21	6.8	7110	6	AX346458	Sequence
9	21	6.8	7110	6	AX349019	Sequence
10	21	6.8	11976	1	AE006249	Lactococc
11	21	6.8	34688	6	AX458543	Sequence
12	21	6.8	10852	9	AC022889	Homo sapi
13	21	6.8	117240	2	AC114561	Mus muscu
14	21	6.8	146814	2	AC022320	Homo sapi
15	21	6.8	161874	9	AC079864	Homo sapi
16	21	6.8	169892	9	AC068397	Homo sapi
17	21	6.8	178756	9	AC104420	Homo sapi
18	21	6.8	184489	2	AC027771	Homo sapi
19	21	6.8	208335	2	AC113206	Mus muscu
20	20	6.4	511	8	BP410032	Brachyste
21	20	6.4	57012	2	AC100122	Mus muscu
22	20	6.4	110000	2	AL353694_2	Continuation (3 of
23	20	6.4	111178	9	AL136129	Human DNA
24	20	6.4	124047	9	AL713966	Human DNA
25	20	6.4	124591	2	AC115457	Rattus no
26	20	6.4	129641	9	AC092904	Homo sapi
27	20	6.4	150546	2	AC130914	Rattus no
28	20	6.4	161200	2	AC102404	Mus muscu
29	20	6.4	172131	2	AC111685	Rattus no
30	20	6.4	191923	2	AC026958	Homo sapi
31	20	6.4	207330	2	AC127055	Rattus no
32	20	6.4	228345	2	AC096226	Rattus no
33	20	6.4	251385	2	AC122377	Mus muscu
34	19	6.1	65	6	AX485140	Sequence
35	19	6.1	335	11	G23898	human STS W
36	19	6.1	1443	9	BC003073	Homo sapi
37	19	6.1	3019	9	AK001383	Homo sapi
38	19	6.1	3476	9	AK092020	Homo sapi
39	19	6.1	4489	9	AB046846	Homo sapi
40	19	6.1	20563	2	AC015202	Drosophil
41	19	6.1	38717	9	AC005357	Homo sapi
42	19	6.1	45870	3	AC024772	Caenorhab
43	19	6.1	52029	9	AL136975	Human DNA
44	19	6.1	55507	2	AC026604	Homo sapi
45	19	6.1	75765	9	AC125490	Homo sapi

ALIGNMENTS

RESULT 1
AX259511
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE

AX259511
Sequence 12 from Patent WO0172775.
AX259511
AX259511.1 GI:16508668
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Chen, S.Y., Hu, P., Recipon, H. and Macina, R.A.
Compositions and methods of diagnosing, monitoring, staging,
imaging and treating lung cancer

AX259511
Sequence 12 from Patent WO0172775.
AX259511
AX259511.1 GI:16508668
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Chen, S.Y., Hu, P., Recipon, H. and Macina, R.A.
Compositions and methods of diagnosing, monitoring, staging,
imaging and treating lung cancer

AX259511
Sequence 12 from Patent WO0172775.
AX259511
AX259511.1 GI:16508668
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Chen, S.Y., Hu, P., Recipon, H. and Macina, R.A.
Compositions and methods of diagnosing, monitoring, staging,
imaging and treating lung cancer

JOURNAL Patent: WO 0172775-A 12 04-OCT-2001;
diaDexus, Inc. (US)
FEATURES Location/Qualifiers
source 1..311
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 107 a 54 c 64 g 86 t
ORIGIN

Query Match 100.0%; Score 311; DB 6; Length 311;
Best Local Similarity 100.0%; Pred. No. 7.7e-164;
Matches 311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTCCAGCTGTGTGCAAGAGATGTGTGGAAATGTCACAGCATCGTATAGCAAGAGC 60
|||||
Db 1 ACTCCAGCTGTGTGCAAGAGATGTGTGGAAATGTCACAGCATCGTATAGCAAGAGC 60
|||||

QY 61 ATATTGGCAACAGCTTGGATGGCCAGCAGAGAGAGCCCAATGTGTGATTCATATTCAC 120
|||||
Db 61 ATATTGGCAACAGCTTGGATGGCCAGCAGAGAGAGCCCAATGTGTGATTCATATTCAC 120
|||||

QY 121 AGTCGAATATTCGAATACATATACACCATATATACATAGACTGTATGTGTGTTCTA 180
|||||
Db 121 AGTCGAATATTCGAATACATATACACCATATATACATAGACTGTATGTGTGTTCTA 180
|||||

QY 181 TACTATAGTGTATGACTTCACTCCATTCAGTGAAGGAGGAGGAGGAGGAGGAGGAGG 240
|||||
Db 181 TACTATAGTGTATGACTTCACTCCATTCAGTGAAGGAGGAGGAGGAGGAGGAGGAGG 240
|||||

QY 241 ATCCATATGGATACAAAAGCAGGATACAAAAGATACATCATCTTGCATTTGC 300
|||||
Db 241 ATCCATATGGATACAAAAGCAGGATACAAAAGATACATCATCTTGCATTTGC 300
|||||

QY 301 AGGTAAGCTT 311
|||||
Db 301 AGGTAAGCTT 311
|||||

RESULT 2
AC090107/c
LOCUS 179396 bp DNA linear HTG 09-MAY-2002
DEFINITION Homo sapiens chromosome 12 clone RP11-643D8, WORKING DRAFT
SEQUENCE, 8 unordered pieces.
AC090107
AC090107.14 GI:20335798
HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens

REFERENCE
AUTHORS Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-Osman, F.R., Allen, C., Alsbrooks, S.L., Anarunge, H.C., Are, J.R., Ayele, M., Banks, T., Barbieri, J., Benton, J., Bimaye, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carton, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy, Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B., Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Louiseged, H.,

Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhiney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Monabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokoko, S., Ogih, M., Okunolu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojubokan, I., Rolfe, M., Ruiz, S., Savary, G., Scherer, S., Scott, G., Shen, H., Shooshari, N., Sisson, I., Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, O., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczyk, R., Woodson, S., Worley, K., Wu, C., Wu, X., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstein, G., and Gibbs, R.

Direct Submission
Unpublished
2 (bases 1 to 179396)
Worley, K.C.
Direct Submission
Submitted (15-FEB-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 179396)
Worley, K.C.
Direct Submission
Submitted (09-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Apr 28, 2002 this sequence version replaced gi:18449697.

Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: HCJT
Center clone name: RP11-643D8
----- Summary Statistics
Sequencing vector: plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 175679 bases at least Q40
Consensus quality: 176511 bases at least Q30
Consensus quality: 176995 bases at least Q20
Estimated insert size: 176936; sum-of-contigs estimation
Quality coverage: 10x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 4459: contig of 4459 bp in length
* 4460 4559: gap of unknown length
* 4560 9953: contig of 5394 bp in length
* 9954 10053: gap of unknown length
* 10054 17744: contig of 7691 bp in length
* 17745 17844: gap of unknown length
* 17845 25913: contig of 8069 bp in length
* 25914 26013: gap of unknown length
* 26014 41674: contig of 15660 bp in length
* 41674 41773: gap of unknown length
* 41774 70908: contig of 29135 bp in length
* 70909 71008: gap of unknown length

* 71009 121798: contig of 50790 bp in length
 * 121799 121898: gap of unknown length
 * 121899 179396: contig of 57498 bp in length.

FEATURES

source
 1. .179396
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="12"
 /clone="RP11-643D8"
 50740 a 38957 c 38965 g 50021 t 713 others

BASE COUNT

ORIGIN

Query Match 50.8%; Score 158; DB 2; Length 179396;
 Best Local Similarity 99.2%; Pred. No. 1.2e-77;
 Matches 258; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 52 GCAAGACATATGGCAACAGCTGGATGCCAGCAGAGAGGCCCAAAATGTGTATTC 111
 Db 138480 GCAAAGACATATGGCAACAGCTGGATGCCAGCAGAGAGGCCCAAAATGTGTATTC 138421

QY 112 ATATTCAGTCTGCAATATTAATGACTACATATACACCATATATACACTAGACTGTATGT 171
 Db 138420 ATATTCAGTCTGCAATATTAATGACTACATATATACACCATATATACACTAGACTGT 138361

QY 172 GTTGTTCTATCTATAGTAGTGTGACTTGAATCCATTCAGTCAAAAAATGGAAGATTA 231
 Db 138360 GTTGTTCTATCTATAGTAGTGTGACTTGAATCCATTCAGTCAAAAAATGGAAGATTA 138301

QY 232 GCTATTTTCATCCATATGGGTACAAAAAGCAGGGTAACAAGAAATCTACATCATCTT 291
 Db 138300 GCTATTTTCATCCATATGGGTACAAAAAGCAGGGTAACAAGAAATCTACATCATCTT 138241

QY 292 GCCATTTCCAGTAAAGCTT 311
 Db 138240 GCCATTTCCAGTAAAGCTT 138221

RESULT 3

LOCUS AX259518 29 bp DNA linear PAT 26-OCT-2001
 DEFINITION Sequence 19 from Patent WO0172775.
 ACCESSION AX259518
 VERSION AX259518.1 GI:16508675

KEYWORDS

synthetic construct.
 synthetic construct.
 artificial sequences.

ORGANISM

1
 Chen, S.Y., Hu, P., Recipon, H. and Macina, R.A.
 Compositions and methods of diagnosing, monitoring, staging,
 imaging and treating lung cancer
 Patent: WO 0172775-A 19 04-OCT-2001;
 diadexus, Inc. (US)

FEATURES

source
 Location/Qualifiers
 1. .29
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="Synthetic"
 11 a 5 c 7 g 6 t

BASE COUNT

ORIGIN
 Query Match 9.3%; Score 29; DB 6; Length 29;
 Best Local Similarity 100.0%; Pred. No. 6.8e-05;
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 AGCATCGTATAGCAAGAGCATATTGGCA 69
 Db 1 AGCATCGTATAGCAAGAGCATATTGGCA 29

RESULT 4

LOCUS AX259516 22 bp DNA linear PAT 26-OCT-2001

DEFINITION Sequence 17 from Patent WO0172775.

ACCESSION AX259516
 VERSION AX259516.1 GI:16508673

KEYWORDS

synthetic construct.
 synthetic construct.
 artificial sequences.

ORGANISM

1
 Chen, S.Y., Hu, P., Recipon, H. and Macina, R.A.
 Compositions and methods of diagnosing, monitoring, staging,
 imaging and treating lung cancer
 Patent: WO 0172775-A 17 04-OCT-2001;
 diadexus, Inc. (US)

FEATURES

source
 Location/Qualifiers
 1. .22
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="Synthetic"
 6 a 3 c 8 g 5 t

BASE COUNT

ORIGIN
 Query Match 7.1%; Score 22; DB 6; Length 22;
 Best Local Similarity 100.0%; Pred. No. 0.6;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 CAAGGAGATGTCTGGGAATGTC 38
 Db 1 CAAGGAGATGTCTGGGAATGTC 22

RESULT 5

LOCUS AF214528S1 4373 bp DNA linear PRI 16-JAN-2000
 DEFINITION Homo sapiens dystrophin (DMD) gene, intron 51, partial sequence.

ACCESSION AF214528
 VERSION AF214528.1 GI:6694738

KEYWORDS

1 of 3
 Homo sapiens.
 Homo sapiens.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 4373)
 Pan, S.Y., Zhang, C. and Liu, Z.L.
 The sequence of intron 51 of the dystrophin gene

TITLE

Unpublished
 2 (bases 1 to 4373)
 Pan, S.Y., Zhang, C. and Liu, Z.L.
 Direct Submission
 Submitted (10-DEC-1999) Dept. of Neurology, Nanfeng Hospital, the
 First Military Medical University, Tonghe, Guangzhou, Guangdong
 510515, P.R.China

JOURNAL

510515, P.R.China

FEATURES

source
 Location/Qualifiers
 1. .4373
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="X"
 /map="Xp21"
 /clone="Pac255E13"
 <1. .>4373
 /gene="DMD"
 /number=51
 1523 a 670 c 918 g 1259 t 3 others

BASE COUNT

ORIGIN
 Query Match 6.8%; Score 21; DB 9; Length 4373;
 Best Local Similarity 100.0%; Pred. No. 2.1;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 206 ATTCAGTCAAAAAAATGGAAG 226
 Db 3707 ATTCAGTCAAAAAAATGGAAG 3727

RESULT 6
AX251243/c
LOCUS AX251243 7110 bp DNA linear PAT 05-OCT-2001
DEFINITION Sequence 211 from Patent WO0168912.
ACCESSION AX251243
VERSION AX251243.1 GI:15984666
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
FEATURES
1 (bases 1 to 7110)
AUTHORS Olek.A., Piepenbrock,C. and Berlin,K.
TITLE Diagnosis of diseases associated with tumor suppressor genes and oncogenes
JOURNAL Patent: WO 0168912-A 211 20-SEP-2001;
Epigenomics AG (DE)
FEATURES
source Location/Qualifiers
1. 7110
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="chemically treated genomic DNA (Homo sapiens)"
BASE COUNT 1749 a 289 c 1822 g 3250 t
ORIGIN
Query Match 6.8%; Score 21; DB 6; Length 7110;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 141 CAATATACACCATATATACTA 161
Db 6679 CAATATACACCATATATACTA 6659
RESULT 7
AX251990/c
LOCUS AX251990 7110 bp DNA linear PAT 05-OCT-2001
DEFINITION Sequence 251 from Patent WO0168911.
ACCESSION AX251990
VERSION AX251990.1 GI:15985345
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
FEATURES
1 (bases 1 to 7110)
AUTHORS Olek.A., Piepenbrock,C. and Berlin,K.
TITLE Diagnosis of diseases associated with the cell cycle
JOURNAL Patent: WO 0168911-A 251 20-SEP-2001;
Epigenomics AG (DE)
FEATURES
source Location/Qualifiers
1. 7110
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="chemically treated genomic DNA (Homo sapiens)"
BASE COUNT 1749 a 289 c 1822 g 3250 t
ORIGIN
Query Match 6.8%; Score 21; DB 6; Length 7110;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 141 CAATATACACCATATATACTA 161
Db 6679 CAATATACACCATATATACTA 6659
RESULT 8
AX346458/c
LOCUS AX346458 7110 bp DNA linear PAT 01-FEB-2002
DEFINITION Sequence 1529 from Patent WO0200928.
ACCESSION AX346458
VERSION AX346458.1 GI:18494344
KEYWORDS

SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1
AUTHORS Olek.A., Piepenbrock,C. and Berlin,K.
TITLE Diagnosis of diseases associated with the immune system
JOURNAL Patent: WO 0200928-A 1529 03-JAN-2002;
Epigenomics AG (DE)
FEATURES
source Location/Qualifiers
1. 7110
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="chemically treated genomic DNA (Homo sapiens)"
BASE COUNT 1749 a 289 c 1822 g 3250 t
ORIGIN
Query Match 6.8%; Score 21; DB 6; Length 7110;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 141 CAATATACACCATATATACTA 161
Db 6679 CAATATACACCATATATACTA 6659
RESULT -9
AX349019/c
LOCUS AX349019 7110 bp DNA linear PAT 06-FEB-2002
DEFINITION Sequence 477 from Patent WO0202807.
ACCESSION AX349019
VERSION AX349019.1 GI:18615054
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
artificial sequences.
REFERENCE 1
AUTHORS Olek.A., Piepenbrock,C. and Berlin,K.
TITLE Diagnosis of diseases associated with cell signalling
JOURNAL Patent: WO 0202807-A 477 10-JAN-2002;
Epigenomics AG (DE)
FEATURES
source Location/Qualifiers
1. 7110
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="chemically treated genomic DNA (Homo sapiens)"
BASE COUNT 1749 a 289 c 1822 g 3250 t
ORIGIN
Query Match 6.8%; Score 21; DB 6; Length 7110;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 141 CAATATACACCATATATACTA 161
Db 6679 CAATATACACCATATATACTA 6659
RESULT 10
AE006249/c
LOCUS AE006249 11976 bp DNA linear BCT 14-MAY-2001
DEFINITION Lactococcus lactis subsp. lactis IL1403 section 11 of 218 of the complete genome.
ACCESSION AE006249 AE005176
VERSION AE006249.1 GI:12722951
KEYWORDS
SOURCE Lactococcus lactis subsp. lactis.
ORGANISM Lactococcus lactis subsp. lactis.
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Lactococcus.
REFERENCE 1 (bases 1 to 11976)
AUTHORS Bolotin,A., Wincker,P., Mauer,S., Jaillon,O., Malarme,K., Weissenbach,J., Ehrlich,S.D. and Sorokin,A.
TITLE The complete genome sequence of the lactic acid bacterium

[illegible]

LKHAVMLKESDKSMIMFPGSRHSAEMKGGVAVIAKMAKVRIVPAVYOGPLTMKGVF
 KROKYSINIGHPTDIDSKIKMDEAGIAEVRNRMEVAFALDKELNDFHYEAR"
 Complement(7941..8432)
 /gene="rmaD"
 /note="L115025"
 complement(7941..8432)
 /gene="rmaD"
 /note="EVIDENCE BY HOMOLOG Y BIO09.08 MarR-family
 regulators. belongs to the MarR family"
 /codon_start=1
 /transl_table=11
 /product="transcriptional regulator"
 /protein_id="AAK04211.1"
 /db_xref="GI:12722959"
 /translation="MTSADLNQOISKELHELHATDNRSQERSWILAOLDNDKLEKEV
 LITSVALHTLSALTEDLTGIELATKLSVTGGVTRAVONLIKYPFLTYQSESDDK
 KIEYHLITVGRKRVATHDKMKIMDLNQLQIFDKYNEKSIILSFSDFNUTEAGLF
 DSE"
 complement(8467..9111)
 /gene="acpD"
 /note="L11551"
 complement(8467..9111)
 /gene="acpD"
 /note="EVIDENCE BY HOMOLOG Y BIO07.01 FATTY ACID AND
 PHOSPHOLIPID METABOLISM. General. converts holo-ACP to
 apo-ACP by hydrolytic cleavage of the phosphopantetheine
 residue from ACP. Probably NAD(P)H dehydrogenase."
 /codon_start=1
 /transl_table=11
 /product="acyl carrier protein phosphodiesterase"
 /protein_id="AAK04212.1"
 /db_xref="GI:12722960"
 /translation="MTLLIILAHPTDDFSWLSATVEEFKKSQESHPLDKIIRDL
 FSKVPALONETFAAKRNKRYAPDTLSADKLLHREVEYLEFLSADKYVFNPNWY
 GFTAEIKQYIDVIAVRKLFERYTENGPIGLGKSLHIQAGGFFYHNEQDPTMAN
 DLGVAIDQTMKMGVLTENRQOLFVEGYARPERADELKEKRAFTSAENEGRAF"
 9285..11882
 /gene="secA"
 /note="L0334"
 9285..11882

Query Match 6.8%; Score 21; DB 1; Length 11976;

Best Local Similarity 100.0%; Pred. No. 2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 213 GAAAAAATGGAGAATTAGC 233

Db 3311 GAAAAAATGGAGAATTAGC 3291

RESULT 11

AX458543/c

LOCUS 34688 bp DNA linear PAT 08-JUL-2002

DEFINITION Sequence 89 from Patent WO0246454.

AX458543

ACCESSION AX458543

VERSION AX458543.1 GI:21725207

KEYWORDS synthetic construct.

SOURCE synthetic construct.

ORGANISM artificial sequences.

REFERENCE Schacht,O.

AUTHORS

TITLE Diagnosis of diseases associated with angiogenesis

JOURNAL Patent: WO 0246454-A 89 13-JUN-2002;

FEATURES Epigenomics AG (DE)

source Location/Qualifiers

1..34688

/organism="synthetic construct"

/db_xref="taxon:32630"

/note="chemically treated genomic DNA (Homo sapiens)"

BASE COUNT 9920 a 444 c 6941 g 17383 t

ORIGIN

Query Match 6.8%; Score 21; DB 6; Length 34688;
 Best Local Similarity 100.0%; Pred. No. 2;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 141 CAATATACCATATATATACTA 161

Db 6679 CAATATACCATATATATACTA 6659

RESULT 12

AC022889

LOCUS

DEFINITION Homo sapiens chromosome 15, clone RP11-492D5, linear PRI 28-APR-2002

AC022889

ACCESSION AC022889

VERSION AC022889.8 GI:20336163

KEYWORDS HTG.

SOURCE Homo sapiens.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 108652)

AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.

TITLE

JOURNAL

REFERENCE 2 (bases 1 to 108652)

AUTHORS

REFERENCE 3 (bases 1 to 108652)

AUTHORS

REFERENCE 4 (bases 1 to 108652)

AUTHORS

REFERENCE 5 (bases 1 to 108652)

AUTHORS

REFERENCE 6 (bases 1 to 108652)

AUTHORS

REFERENCE 7 (bases 1 to 108652)

AUTHORS

REFERENCE 8 (bases 1 to 108652)

AUTHORS

REFERENCE 9 (bases 1 to 108652)

AUTHORS

REFERENCE 10 (bases 1 to 108652)

AUTHORS

REFERENCE 11 (bases 1 to 108652)

AUTHORS

REFERENCE 12 (bases 1 to 108652)

AUTHORS

REFERENCE 13 (bases 1 to 108652)

AUTHORS

REFERENCE 14 (bases 1 to 108652)

AUTHORS

REFERENCE 15 (bases 1 to 108652)

AUTHORS

REFERENCE 16 (bases 1 to 108652)

AUTHORS

REFERENCE 17 (bases 1 to 108652)

AUTHORS

REFERENCE 18 (bases 1 to 108652)

AUTHORS

REFERENCE 19 (bases 1 to 108652)

AUTHORS

REFERENCE 20 (bases 1 to 108652)

AUTHORS

REFERENCE 21 (bases 1 to 108652)

AUTHORS

REFERENCE 22 (bases 1 to 108652)

AUTHORS

REFERENCE 23 (bases 1 to 108652)

AUTHORS

REFERENCE 24 (bases 1 to 108652)

AUTHORS

REFERENCE 25 (bases 1 to 108652)

AUTHORS

REFERENCE 26 (bases 1 to 108652)

AUTHORS

REFERENCE 27 (bases 1 to 108652)

AUTHORS

Submitted (15-APR-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 4 (bases 1 to 108652)

Direct Submission
 TITLE
 JOURNAL
 AUTHORS

Submitted (15-APR-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 4 (bases 1 to 108652)

Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L.,
 Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,
 Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,
 Cook, A., Cooke, P., DeKrellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
 Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
 Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
 Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Kanat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R.,
 Landers, T., Lehoczyk, J., Levine, R., Lindblad-Toh, K., Liu, G.,
 MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,
 McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Meneus, L.,
 Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,
 Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
 Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
 Raymond, C., Retta, R., Riback, M., Riley, R., Rise, C., Rogov, P.,
 Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnupack, R.,
 Seaman, S., Severi, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
 Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
 Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE

Submitted (28-APR-2002) Whitehead Institute/MIT Center for Genome

JOURNAL

Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT

On Apr 28, 2002 this sequence version replaced gi:20149479.
 All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www.seq.wi.mit.edu

Contact: sequence.submissions@genome.wi.mit.edu

----- Project Information

Center project name: L6227

Center clone name: 492_D_5

Only the first 108.7 kilobases of this clone are being submitted.
 The remainder overlaps accession number AC068397 [WICGR project
 L10279].

FEATURES

source

1. .108652
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="15"
 /map="15"
 /clone="RP11-492B5"
 /clone_lib="RPC1-11 Human Male BAC"
 complement(55. .104)
 /rpt_family="MIR"
 complement(144. .576)
 /rpt_family="LTR54B"
 complement(1235. .1316)
 /rpt_family="LTR50"
 1317. .1613
 /rpt_family="AluSx"
 complement(1614. .1769)
 /rpt_family="LTR50"
 2365. .3030
 /rpt_family="L1M4"
 3101. .3280
 /rpt_family="L3"
 3739. .3764
 /rpt_family="AT_rich"
 complement(4174. .4299)
 /rpt_family="MIR3"
 complement(5049. .5339)
 /rpt_family="AluJo"
 6164. .6184
 /rpt_family="AT_rich"
 complement(6620. .7051)
 /rpt_family="LTR41"

repeat_region complement(7041. .7228)
 /rpt_family="LTR41"
 repeat_region complement(7367. .7516)
 /rpt_family="L1MC/D"
 repeat_region complement(8905. .9307)
 /rpt_family="MSTA"
 repeat_region complement(9420. .9790)
 /rpt_family="THE1B"
 9854. .9981
 /rpt_family="L1MCC"
 complement(9982. .10341)
 /rpt_family="THE1B"
 10342. .11566
 /rpt_family="L1MCC"
 11647. .11666
 /rpt_family="(TTAA)n"
 complement(11667. .11852)
 /rpt_family="AluSg"
 11926. .12144
 /rpt_family="L1MCI"
 12145. .12456
 /rpt_family="AluSx"
 complement(12315. .12320)
 /note="<30 qual SNGL region"
 12457. .12934
 /rpt_family="L1MCI"
 12932. .13296
 /rpt_family="L1MCC/D"
 13295. .13703
 /rpt_family="L1MCC"
 complement(13735. .13804)
 /rpt_family="U6"
 13809. .13971
 /rpt_family="L1MCC"
 13995. .14034
 /rpt_family="AT_rich"
 complement(14295. .14452)
 /rpt_family="MER58A"
 14915. .14935
 /rpt_family="AT_rich"
 complement(15013. .15242)
 /rpt_family="MER58A"
 complement(15771. .16240)
 /rpt_family="L1MB4A"
 16315. .16590
 /rpt_family="AluJb"
 16704. .16946
 /rpt_family="MER8"
 17267. .17294
 /rpt_family="AT_rich"
 17623. .17643
 /rpt_family="AT_rich"
 18343. .18453
 /rpt_family="MIR3"
 complement(18441. .18445)
 /note="<30 qual SNGL region"
 18680. .18797
 /rpt_family="MIR"
 complement(19470. .19724)
 /rpt_family="L2"
 20186. .20206
 /rpt_family="AT_rich"
 20647. .20670
 /rpt_family="AT_rich"
 complement(20785. .20918)
 repeat_region

Query Match 6.8%; Score 21; DB 9; Length 108652;

Best Local Similarity 100.0%; Pred. No. 2;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 277 AATCATCATCATTCGCATT 297

|||||

Db 60587 AATCATCATCATTCGCATT 60607

http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L23272
Center clone name: 54_F.12
----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 115177 bases at least Q40
Consensus quality: 115810 bases at least Q30
Consensus quality: 116125 bases at least Q20
Insert size: 115000; agarose-fp
Quality coverage: 7.2 in Q20 bases; agarose-fp
Quality coverage: 7.1 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 35824: contig of 35824 bp in length
* 35825 35924: gap of 100 bp
* 35925 36989: contig of 1065 bp in length
* 36990 37089: gap of 100 bp
* 37090 93393: contig of 56304 bp in length
* 93394 93493: gap of 100 bp
* 97874 97973: contig of 4380 bp in length
* 97974 115048: contig of 17075 bp in length
* 115049 115148: gap of 100 bp
* 115149 117240: contig of 2092 bp in length.
----- Location/Qualifiers
source
1. 117240
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="RP23-54F12"
/clone_lib="RPCI-23 Female Mouse BAC"
misc_feature
1. 35824
/note="assembly_fragment"
clone_end:SP6
vector_side:left
35925..36989
misc_feature
/note="assembly_fragment"
37090..93393
/note="assembly_fragment"
93494..97873
misc_feature
/note="assembly_fragment"
97974..115048
/note="assembly_fragment"
115149..117240
/note="assembly_fragment"
clone_end:T7
vector_side:right
BASE COUNT 36640 a 21784 c 21239 g 37053 t 524 others
ORIGIN
Query Match 6.8%; Score 21; DB 2; Length 117240;
Best Local Similarity 100.0%; Pred. NO. 2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 172 GTGTCTTACTACTAGTGCAT 192
|||||
DB 17902 GTGTCTTACTACTAGTGCAT 17922

RESULT 13
AC114561
LOCUS
DEFINITION
AC114561.3 GI:21360035
VERSION
KEYWORDS
SOURCE
ORGANISM
house mouse.
Mus musculus
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 117240)
Birren, B., Linton, L., Nusbaum, C., Lander, E., All, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,
Boukhalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,
Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,
Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Gage, D., Gage, D.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Galagan, J., Gardyna, S.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R.,
Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G.,
Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,
McCarthy, M., McEwan, P., McKernan, K., Meldrim, J., Meneus, L.,
Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,
Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R.,
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (10-MAR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 117240)
Birren, B., Linton, L., Nusbaum, C., Lander, E., All, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,
Boukhalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,
Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,
Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Gage, D., Gage, D.,
Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L.,
Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I.,
Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., LaRocque, K.,
Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K.,
Liu, G., Maclean, C., Macdonald, P., Major, J., Marquis, N.,
Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrim, J.,
Meneus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C.,
Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,
Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C.,
Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S.,
Schuback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N.,
Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S.,
Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,
Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (09-JUN-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 9, 2002 this sequence version replaced gi:21327469.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)


```
repeat_region /standard_name="152451"
repeat_region complement(18180..18685)
repeat_region /rpt_family="MLT1D"
repeat_region 20533..20581
repeat_region /rpt_family="AT_rich"
repeat_region 20837..20907
repeat_region /rpt_family="(CAGA)n"
repeat_region 21387..21438
repeat_region /rpt_family="AT_rich"
STS 22820..22968
STS /standard_name="71574"
STS 22821..22948
STS /standard_name="88914"
repeat_region 22875..22917
repeat_region /rpt_family="(CA)n"
repeat_region 22950..23106
repeat_region /rpt_family="MIR"
repeat_region 23744..23853
repeat_region /rpt_family="LTR16A"
repeat_region complement(23854..24205)
repeat_region /rpt_family="MLT1A1"
repeat_region 24206..24526
repeat_region /rpt_family="LTR16A"
repeat_region 26053..26088
repeat_region /rpt_family="AT_rich"
repeat_region 26406..26547
repeat_region /rpt_family="MER103"
repeat_region complement(26702..26757)
repeat_region /rpt_family="LIMA10"
repeat_region 26758..27120
repeat_region /rpt_family="MER7A"
repeat_region complement(27121..27343)
repeat_region /rpt_family="LIMA10"
repeat_region complement(27354..27413)
repeat_region /rpt_family="AluY"
repeat_region complement(27414..27595)
repeat_region /rpt_family="AluY"
repeat_region 28481..28559
repeat_region /rpt_family="MLT1L"
```

```
Query Match 6.8%; Score 21; DB 9; Length 161874;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 206 ATTCAGTGAATAAATGGAAG 226
Db 35393 ATTCAGTGAATAAATGGAAG 35373
|||||
```

```
Search completed: July 7, 2003, 19:41:25
Job time : 1139 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 7, 2003, 18:14:17 ; Search time 217 Seconds
(without alignments)
3227.518 Million cell updates/sec

Title: US-09-820-425b-12

Perfect score: 311

Sequence: 1 actccagctctgtgtcaag.....gccattgcaggtaaagctt 311

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_101002.*

```

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	311	100.0	311	22	AA19594 Human Lung cancer
2	41.4	13.3	1768	21	AA19594 Human secreted pro
3	41.4	13.3	1831	22	AA19594 Human CDNA sequenc
4	40.4	13.0	6292	22	AA19594 Tumour suppressor
5	38.2	12.3	1989	22	AA19594 Mouse glycosyl sul
6	38.2	12.3	32392	24	ABL56203 AMEPV genome fragm
7	36.8	11.8	1851	20	AA19594 Nucleic acid sequ
8	36.8	11.8	1887	20	AA19594 Nucleic acid sequ
9	36.8	11.8	19236	24	ABL56203 Human chemically m

10	36.8	11.8	1082138	21	AA19594 Arabidopsis thalia
11	36.4	11.7	475	22	AA19594 Human foetal liver
12	36.4	11.7	475	22	AA19594 Human brain expres
13	36.4	11.7	475	22	AA19594 Human bone marrow
14	36.4	11.7	475	22	AA19594 Probe #6751 used t
15	36.4	11.7	475	24	AA19594 Human genome-deriv
16	36.4	11.7	598	22	AA19594 Human foetal liver
17	36.4	11.7	598	22	AA19594 Human bone marrow
18	36.4	11.7	598	22	AA19594 Probe #8083 used t
19	36.2	11.6	548	22	AA19594 Human immune/haema
20	36.2	11.6	16688	24	ABL32321 Human immune syste
21	36	11.6	2145	21	AA19594 Human secreted pro
22	35.8	11.5	6341	24	ABL32321 Human immune syste
23	35.8	11.5	61020	22	AA19594 Tumour suppressor
24	35.6	11.4	1025	22	AA19594 Human SCN3A genomi
25	35.4	11.4	19345	24	ABL56203 Human chemically m
26	35.2	11.3	9502	22	AA19594 Tumour suppressor
27	35.2	11.3	17918	24	AA19594 Human gene regulat
28	35	11.3	6707	22	AA19594 Tumour suppressor
29	35	11.3	72604	20	AA19594 Genomic sequence o
30	35	11.3	72604	24	ABL32321 Human HKMG1 gene.
31	34.8	11.2	4779	23	ABL27436 Drosophila melanog
32	34.8	11.2	10078	23	ABL19228 Human chromosome 1
33	34.6	11.1	376	22	AA19594 Human polynucleoti
34	34.6	11.1	6317	24	ABL32408 Human immune syste
35	34.6	11.1	125910	21	AA19594 Human KCNQ5 (KCN6q
36	34.6	11.1	172637	24	ABL32408 Human voltage-acti
37	34.6	11.1	1863	16	AA19594 Ovalbumin pl-yoval
38	34.4	11.1	3425	22	AA19594 Human digestive sy
39	34.4	11.1	3425	22	AA19594 Human liver associ
40	34.4	11.1	3425	22	AA19594 Human liver associ
41	34.4	11.1	3425	24	ABL18526 Drosophila melanog
42	34.4	11.1	5256	23	ABL18526 Drosophila melanog
43	34.4	11.1	6300	23	ABL29704 Human digestive sy
44	34.4	11.1	6945	22	AA19594 Human digestive sy
45	34.4	11.1	6945	22	AA19594 Human digestive sy

ALIGNMENTS

RESULT 1

AA19594

ID AA19594 standard; cDNA; 311 BP.

XX AA19594;

XX AA19594;

XX 18-DEC-2001 (first entry)

DT Human Lung cancer Specific Gene (LSG) Sg1ng020 cDNA.

XX Human; Lung cancer Specific Gene; LSG; cytostatic; vaccine; gene therapy;

XX cancer; immune response; cell disorder; tissue; ss.

OS Homo sapiens.

PN WO200172775-A2.

PD 04-OCT-2001.

XX 29-MAR-2001; 2001WO-US10051.

XX 29-MAR-2000; 2000US-192921P.

XX (DIAD-) DIADEXUS INC.

XX Chen S, Hu P, Recipon H, Macina RA;

XX WPI; 2001-616467/71.

XX New lung cancer specific genes and polypeptides encoded by the genes,

XX useful for detecting, diagnosing, monitoring, staging, prognosticating,

XX imaging and treating lung cancer

XX Claim 1; Page 74; 76pp; English.

XX The present sequence is lung cancer specific genes (LSG) cDNA. LSG and

XX its protein are useful for diagnosing, detecting, monitoring, staging,

XX prognosticating, imaging and treating lung cancer in a patient by

XX determining the levels of LSG in cells, tissues or bodily fluids in a

XX patient and comparing the determined levels of LSG with levels of LSG

XX in cells, tissues or bodily fluids from a normal human control, where a

XX change in determined levels of LSG in the patient versus normal control

XX is associated with the presence of LSG in the patient and treating

XX identifying potential therapeutic agents for use in imaging and treating

XX lung cancer. LSG antibody labelled with paramagnetic ions or a

XX radioisotope is useful for imaging lung cancer in a patient. LSG antibody

XX conjugated to a cytotoxic agent is useful for treating lung cancer in a

XX patient. LSG vaccine is useful for inducing an immune response against a

XX LSG protein and for treating lung cancer in a patient. LSG and its

XX protein are useful as diagnostic markers for lung cancer and for

XX diagnosis and treatment of disorders of cells, tissues and organisms.

XX SQ Sequence 311 BP; 107 A; 54 C; 64 G; 86 T; 0 other;

Query Match 100.0%; Score 311; DB 22; Length 311;

Best Local Similarity 100.0%; Pred. No. 5.7e-78;

Matches 311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTCCAGCTCTGTGTCAGGAGATGTGCTGGATGTGCAGCATCGTATAGCAAGAGC 60

DB 1 ACTCCAGCTCTGTGTCAGGAGATGTGCTGGATGTGCAGCATCGTATAGCAAGAGC 60

QY 61 ATATTGGCAACAGCTTGGATGGCCAGCAGAGAGGCCCAATGTGTATTCATTCAC 120

DB 61 ATATTGGCAACAGCTTGGATGGCCAGCAGAGAGGCCCAATGTGTATTCATTCAC 120

QY 121 AGTCGATATTAAGTAACTACATATACATATACCATATATAGTGTGTGTCTTA 180

DB 121 AGTCGATATTAAGTAACTACATATACATATATAGTGTGTGTGTCTTA 180

QY 181 TACTATAGTGATGACTTGACCTCCATTGAGTGAAGGAGGAGGAGGAGGAGGAGG 240

DB 181 TACTATAGTGATGACTTGACCTCCATTGAGTGAAGGAGGAGGAGGAGGAGGAGG 240

QY 241 ATCCATATGGGATACAAAAAGCAGGGGTACAAAAAGAAATCTACATCATCTTGC 300

DB 241 ATCCATATGGGATACAAAAAGCAGGGGTACAAAAAGAAATCTACATCATCTTGC 300

QY 301 AGGTAAGCTT 311

DB 301 AGGTAAGCTT 311

RESULT 2

AAC96919/C

ID AAC96919 standard; cDNA; 1768 BP.

XX AC AAC96919;

XX 22-FEB-2001 (first entry)

XX Human secreted protein cDNA #20.

XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;

XX antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;

XX vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;

XX cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;

XX neurological disease; infection; human; secreted protein; ss.

XX Homo sapiens.

XX WO200061624-A1.

XX 19-OCT-2000.

XX

PF 06-APR-2000; 2000WO-US08980.

XX 09-APR-1999; 99US-0128700.

PR 20-JAN-2000; 2000US-0176930.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM, Komatsoulis G;

XX WPI; 2000-656324/63.

DR P-PSDB; AAB521122.

XX New nucleic acid molecules encoding human secreted proteins, used in

XX preventing, treating or ameliorating a disorder, e.g. Alzheimer's and

XX Parkinson's diseases and cancers

XX Claim 1; Page 406; 478pp; English.

XX The invention relates to the isolation of genes AAC96900-C96947 encoding

XX 48 human secreted proteins AAB52104-B52150. The genes can be used to

XX generate fusion proteins by linking to the gene for the human

XX immunoglobulin G Fc portion (SEQID) for increasing the stability of

XX the fusion protein as compared to the human protein only. The genes and

XX proteins are useful for preventing, ameliorating or treating medical

XX conditions, e.g. by protein or gene therapy. The genes are isolated

XX from a range of human tissues disclosed in the specification. The

XX nucleic acids, proteins, antibodies and (ant)agonists are useful in

XX the diagnosis, treatment and prevention of: (a) cancer, e.g. breast

XX and ovarian cancer, and other cancers of the adrenal gland, bone, bone

XX marrow, breast, gastrointestinal tract, liver, lung, or urogenital;

XX (b) immune disorders e.g. Addison's disease, allergies, autoimmune

XX haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's

XX disease, multiple sclerosis, rheumatoid arthritis and ulcerative

XX colitis; (c) cardiovascular disorders such as myocardial ischaemias; (d)

XX wound healing; (e) neurological diseases e.g. cerebral anoxia and

XX epilepsy; and (f) infectious diseases such as viral, bacterial, fungal

XX and parasitic infections.

XX SQ Sequence 1768 BP; 497 A; 310 C; 334 G; 627 T; 0 other;

Query Match 13.3%; Score 41.4; DB 21; Length 1768;

Best Local Similarity 51.3%; Pred. No. 0.077;

Matches 96; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 109 TTCATATTCACATAGTCGAATAATTGAATCTACTACATATACCATATATAGTACTGTA 168

DB 1169 TCGAAAGCAGTACACCAAGAAATAATAGGGGAAAAAATATGCTTCTTAACAGTC 1110

QY 169 TGTGTTGTTCTATATAGTATGATTCAGTTCAGTCACTCACTCACTCACTCACTCACT 228

DB 1109 TTTCCCTTTCTGAAGTAAAGTCTTTAAAAAAAATCAACAGTACAAAAATATTATC 1050

QY 229 TTAGCTATTGTTGATCCATATGGATACAAAAAGAGGAGGTAACAAAGATCTACATCAT 288

DB 1049 ATATATTTTAAAAACCTATCATCAAGTCCAAATCTCTGTAAAAAAGAAGATGTTTCTT 990

QY 289 CTGCGCA 295

DB 989 CAAGTCA 983

RESULT 3

AAH14247/C

ID AAH14247 standard; cDNA; 1831 BP.

XX AAH14247;

XX 26-JUN-2001 (first entry)

XX Human cDNA sequence SEQ ID NO:11551.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX

```
OS Homo sapiens.
XX EP1074617-A2.
XX 07-FEB-2001.
XX 28-JUL-2000; 2000EP-0116126.
XX 29-JUL-1999; 99JP-0248036.
XX 27-AUG-1999; 99JP-0300253.
XX 11-JAN-2000; 2000JP-0118776.
XX 02-MAY-2000; 2000JP-0183767.
XX 09-JUN-2000; 2000JP-0241899.
XX (HELI-) HELIX RES INST.
XX Ota T, Isoqai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602
XX full-length cDNAs defined in the specification and for the detection
XX and/or diagnosis of the abnormality of the proteins encoded by the
XX full-length cDNAs -
XX
XX Claim 8; SEQ ID 11551; 2537pp + CD ROM; English.
XX
XX The present invention describes primer sets for synthesising 5602
XX full-length cDNAs defined in the specification. Where a primer set
XX comprises: (a) an oligo-dT primer and an oligonucleotide complementary
XX to the complementary strand of a polynucleotide which comprises one of
XX the 5602 nucleotide sequences defined in the specification, where the
XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX of an oligonucleotide comprising a sequence complementary to the
XX complementary strand of a polynucleotide which comprises a 5'-end
XX sequence and an oligonucleotide comprising a sequence complementary to a
XX polynucleotide which comprises a 3'-end sequence, where the
XX oligonucleotide comprises at least 15 nucleotides and the combination of
XX the 5'-end sequence/3'-end sequence is selected from those defined in
XX the specification. The primer sets can be used in antisense therapy and
XX in gene therapy. The primers are useful for synthesising polynucleotides,
XX particularly full-length cDNAs. The primers are also useful for the
XX detection and/or diagnosis of the abnormality of the proteins encoded by
XX the full-length cDNAs. The primers allow obtaining of the full-length
XX cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
XX AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
XX AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
XX represent oligonucleotides, all of which are used in the exemplification
XX of the present invention.
XX
XX Sequence 1831 BP; 468 A; 360 C; 401 G; 602 T; 0 other;
XX
XX Query Match 13.3%; Score 41.4; DB 22; Length 1831;
XX Best Local Similarity 51.3%; Pred. No. 0.078; Indels 0; Gaps 0;
XX Matches 96; Conservative 0; Mismatches 91;
XX
XX QY 109 TTCAATTCACCTAGTCGATATTAATGAATACATATACATATACATATACATGTA 168
XX Db 1220 TCGAAAAGCAGTAACACAGAAATAAATAGGCAAAATATCGATCTTTCTTAACAGTC 1161
XX
XX QY 169 TGTGTGTTCCTATATAGTCATTCGACTTGAACCTCCATTCAGTGAAGAAATGGAAGAA 228
XX Db 1160 TTTCCTCTCTCTGAAGTAAAGCTTTTAAAAAAAATCAACAAGTACAAAAATATTATTC 1101
XX
XX QY 229 TTAGCTATTGTATCATATGGATACAAAAGCAGGTAACAAAAGATCTACATCAT 288
XX Db 1100 ATATATTTTAAACCTATCATCAAGTCCAAATCTCTGTAAGAAAGAAATGTTTCTT 1041
XX
XX 289 CTTGCCA 295
XX 1040 CAAGTCA 1034
```

```
RESULT 4
AAS46736/C
ID AAS46736 standard; DNA; 6292 BP.
XX
XX AAS46736;
AC
XX 18-DEC-2001 (first entry)
DT
XX Tumour suppressor gene derived chemically modified sequence #460.
XX
XX Human; tumour suppressor gene; oncogene; antitumour; cytostatic;
XX cancer; tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;
XX cytosine methylation; ds.
OS Homo sapiens.
XX
XX WO200168912-A2.
PN
XX 20-SEP-2001.
PD
XX
XX 15-MAR-2001; 2001WO-EP02955.
PF
XX
XX 15-MAR-2000; 2000DE-1013847.
PR
XX 06-APR-2000; 2000DE-1019058.
PR
XX 07-APR-2000; 2000DE-1019173.
PR
XX 30-JUN-2000; 2000DE-1032529.
PR
XX 01-SEP-2000; 2000DE-1043826.
PR
XX (EPIG-) EPIGENOMICS AG.
PA
XX Olek A, Piepenbrock C, Berlin K;
PI
XX WPI; 2001-602752/68.
PT
XX
XX Fragments of chemically modified genes associated with tumour suppressor
XX genes and oncogenes, useful in designing primers and probes for
XX analysing diseases associated with cytosine methylation state e.g.
XX cancer -
XX
XX Claim 1; SEQ ID No 460; 27pp; English.
XX
XX The invention relates to a nucleic acid comprising a sequence of 18
XX bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with
XX bisulphite, of genes associated with tumour suppression and
XX oncogenes having a sequence taken from 536 (actually 533 since
XX numbers 408, 458 and 500 are missing from the sequence listing) sequences
XX (Ss) and sequences complementary to (Ss). The nucleic acid may be a
XX peptide nucleic acid-oligomer (PNA) of at least 9 nucleotides and may
XX form part of a set of probes for detecting the cytosine methylation state
XX and/or single nucleotide polymorphisms and also to be used in an
XX array for analysing diseases associated with CpG dinucleotides e.g.
XX cancers and tumours. The probes can also be used in a method for
XX ascertaining genetic and/or epigenetic parameters for the diagnosis
XX and/or therapy of existing diseases or the predisposition to specific
XX diseases, by analysing cytosine methylations. The parameters may be
XX compared to another set of genetic and/or epigenetic parameters, the
XX differences serving as basis for diagnosis and/or prognosis events which
XX are disadvantageous to patients. The present sequence is one of the
XX 533 genomic sequences derived from tumour suppressor genes and
XX oncogenes. Sequences with even numbered Seq ID numbers are the
XX complementary sequence of the corresponding odd numbered sequence (e.g.
XX ID 2 and ID1, ID 536 and ID 535, except for those whose partner sequence
XX is missing).
XX
XX Note: The sequence data for this patent did not form part
XX of the printed specification, but was obtained in electronic
XX format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 6292 BP; 1642 A; 99 C; 1610 G; 2941 T; 0 other;
XX
XX Query Match 13.0%; Score 40.4; DB 22; Length 6292;
XX Best Local Similarity 52.4%; Pred. No. 0.22;
```


PS Disclosure; Page 226-242; 326pp; English.

XX The invention relates to a recombinant entomopox virus (EPV) vector,
XX comprising a polynucleotide encoding a protein operably linked with a
CC heterologous promoter sequence. The invention also concerns methods for
CC providing gene therapy for genetic deficiency disorders. Vectors of the
CC invention are useful for delivering a polynucleotide encoding a protein
CC to a vertebrate cell preferably a mammalian cell, such as a human cell.
CC The vector is introduced into the vertebrate cell by infection in a viral
CC particle, or by transfection, transduction, or injection either in vitro
CC or in vivo. The vector is useful for the delivery and expression of
CC biologically useful proteins in gene therapy protocols, and for
CC delivering large DNA segments for engineering of vertebrate cells.
CC Polynucleotides of the invention have applications in techniques such as
CC their use as insertion sites for foreign genes of interest, hybridisation
CC probes, for chromosome and gene mapping, in PCR technologies, and in the
CC production of sense or antisense nucleic acids. Vectors of the invention
CC provide for stable integration and expression of heterologous DNA in host
CC cells, and are adapted for accepting large heterologous polynucleotide
CC inserts which can be delivered in an infected or transformed cell and
CC expressed in a stable fraction. The current sequence represents a
CC fragment of the genome of the genus B entomopoxvirus from amsacta moorei
CC (AmEPV).

XX
SQ Sequence 32392 BP; 13748 A; 2577 C; 2550 G; 13517 T; 0 other;

Query Match 12.3%; Score 38.2; DB 24; Length 32392;
Best Local Similarity 58.7%; Pred. No. 1.5;
Matches 84; Conservative 0; Mismatches 58; Indels 1; Gaps 1;

QY 135 ATACACAAATACACCAATATATACAGCTGTATGTTGTTCTATCTATAGTGATTG 194
DB 8244 ATTATATATATATAAATGGATACCTTTACCATCGCAATTATTATTTAAATATTATA 8303
QY 195 ACTTCACTCCATTCAGTGAAATAATGAAGAATAGCTATTGTTATCCATATGGATA 254
DB 8304 ATTTAGATATATGTTATATATATATTTGTATAA-TATTGATTTTATACAGATGATA 8362
QY 255 CAAAAAGCAGGGTAACAAAGA 277
DB 8363 TATAAATAAATAAATAAATAA 8385

RESULT 7
AAAX99551/c
ID AAX99551 standard; DNA; 1851 BP.

XX
AC AAX99551;
XX
DT 05-OCT-1999 (first entry)
XX
DE Nucleic acid sequence from U. urealyticum.
XX
KW Ureaplasma urealyticum; nucleic acid detection; infection; pathogen;
KW human urogenital tract; pregnancy; neonatal disease; drug therapy;
KW suppurative arthritis; ss.

XX
OS Ureaplasma urealyticum.
XX
PN WO9939007-Al.
XX
PD 05-AUG-1999.
XX
PF 29-JAN-1999; 99WO-US01972.
XX
PR 30-JAN-1998; 98US-0073189.
XX
PA (UABR-) UAB RES FOUND.
XX
PI Cassell GH, Chen EY, Glass JI, Glass JS, Heiner CR;
PI Lefkowitz E;
XX
DR WPI; 1999-469343/39.

XX Detection of Ureaplasma urealyticum using novel genes, probes and
PT primers
XX
PS Claim 1; Page 38; 110pp; English.

XX The present invention provides methods for the detection and diagnosis
CC of Ureaplasma urealyticum infection. It provides novel genes
CC (AAX99501-681) that can be used as a source of primers and probes for the
CC detection and/or quantification of U. urealyticum in a biological
CC sample. The probes that can be used in the method of the invention by
CC forming target:probe complex is complementary to a region selected from
CC one of the 181 nucleic acid sequences (AAX99501-681). U. urealyticum is
CC an opportunistic pathogen of the human urogenital tract that is a
CC significant cause of adverse pregnancy outcome, neonatal disease, and
CC suppurative arthritis. As the infections are commonly asymptomatic, it is
CC important to have specific and sensitive methods for detecting their
CC presence in a patient. Also, as the pathogen has no current antibiotic
CC directed specifically against it, it would be advantageous to isolate and
CC detect gene sequences which are unique to it, and utilise these as a
CC basis for diagnosis of U. urealyticum infection as well as to develop new
CC and improved drug therapies. The present invention provides such novel
CC polynucleotide sequences (AAX99501-681).

XX Sequence 1851 BP; 587 A; 187 C; 203 G; 874 T; 0 other;

Query Match 11.8%; Score 36.8; DB 20; Length 1851;
Best Local Similarity 50.6%; Pred. No. 1.5;
Matches 89; Conservative 0; Mismatches 87; Indels 0; Gaps 0;
QY 117 CACTAGTTCGAATAATTGAATCTACAATATATACACCATATATAGACTGTATGTTGT 176
DB 1618 CATTATAAATTTATCAGAGCAAAATTTAAACACAGCAACAATTTAAAAATTTAAT 1559
QY 177 TCTATCTACTAGTGTGATTCAGTTCAGTTCATTCAGTGAAGAAATGGAAGATTTAGCTAT 236
DB 1558 TACACAAGTGTGTTTAAATGAATTAGATTCAATCAAAAACATGAGTGTATTATAGA 1499
QY 237 TTGTATCCATATGGGATACAAAAAGCAGGGTAACAAAAGATCTACATCATCTTG 292
DB 1498 TCTTGTGTAATTAGAAACACTAAAAACAAATTAACAAGATTTTAAATCAATATG 1443

RESULT 8
AAAX99539/c
ID AAX99539 standard; DNA; 1887 BP.

XX
AC AAX99539;
XX
DT 05-OCT-1999 (first entry)
XX
DE Nucleic acid sequence from U. urealyticum.
XX
KW Ureaplasma urealyticum; nucleic acid detection; infection; pathogen;
KW human urogenital tract; pregnancy; neonatal disease; drug therapy;
KW suppurative arthritis; ss.

XX
OS Ureaplasma urealyticum.
XX
PN WO9939007-Al.
XX
PD 05-AUG-1999.
XX
PF 29-JAN-1999; 99WO-US01972.
XX
PR 30-JAN-1998; 98US-0073189.
XX
PA (UABR-) UAB RES FOUND.
XX
PI Cassell GH, Chen EY, Glass JI, Glass JS, Heiner CR;
PI Lefkowitz E;
XX
DR WPI; 1999-469343/39.

PT brains -
XX Example 4; SEQ ID NO: 6520; 650pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is one of the probes of the
CC invention.
XX
SQ Sequence 475 BP; 134 A; 87 C; 86 G; 168 T; 0 other;
Query Match 11.7%; Score 36.4; DB 22; Length 475;
Best Local Similarity 52.7%; Pred. No. 1.3;
Matches 79; Conservative 0; Mismatches 71; Indels 0; Gaps 0;
QY 101 ATGTGATTCATATTCACACTAGTCGAAATAATTGAATCTACAATATACACATATATACT 160
DB 53 ATGTACTTTTATATTAATTTTAAATATTTATGTATACACTACTATACAAATATATAGAT 112
QY 161 AGACTGTATGTGTTCTATATAGTATGATTCAGTTCGAACTTGAATCCATTCAGTGAATAA 220
DB 113 ATACAGTAATGATTAATGCAATCTATTTTAAATCTATAGATCCAGAAATGTGGTAAACT 172
QY 221 TGAAGAATTAAGTATTTGTCATCCATATGG 250
DB 173 GAGCACTAATAGAGATTTAGTAATCTGG 202
RESULT 13
AAK32214
ID AAK32214 standard; DNA: 475 BP.
XX
AC AAK32214;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed single exon probe SEQ ID NO: 6771.
XX
DE Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS Homo sapiens.
XX
PN WO200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00668.
XX
PR 04-FEB-2000; 2000US-0180312.
XX
PR 26-MAY-2000; 2000US-0207456.
XX
PR 30-JUN-2000; 2000US-0608408.
XX
PR 03-AUG-2000; 2000US-0632366.
XX
PR 21-SEP-2000; 2000US-0234687.
XX
PR 27-SEP-2000; 2000US-0236359.
XX
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488900/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human bone marrow -
XX
XX Example 4; SEQ ID NO: 6771; 658pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human

CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC the probes of the invention.
XX
SQ Sequence 475 BP; 134 A; 87 C; 86 G; 168 T; 0 other;
Query Match 11.7%; Score 36.4; DB 22; Length 475;
Best Local Similarity 52.7%; Pred. No. 1.3;
Matches 79; Conservative 0; Mismatches 71; Indels 0; Gaps 0;
QY 101 ATGTGATTCATATTCACACTAGTCGAAATAATTGAATCTACAATATACACATATATACT 160
DB 53 ATGTACTTTTATATTAATTTTAAATATTTATGTATACACTACTATACAAATATATAGAT 112
QY 161 AGACTGTATGTGTTCTATATAGTATGATTCAGTTCGAACTTGAATCCATTCAGTGAATAA 220
DB 113 ATACAGTAATGATTAATGCAATCTATTTTAAATCTATAGATCCAGAAATGTGGTAAACT 172
QY 221 TGAAGAATTAAGTATTTGTCATCCATATGG 250
DB 173 GAGCACTAATAGAGATTTAGTAATCTGG 202
RESULT 14
AAI38065
ID AAI38065 standard; DNA: 475 BP.
XX
AC AAI38065;
XX
DT 17-OCT-2001 (first entry)
XX
DE Probe #6751 used to measure gene expression in human placenta sample.
XX
DE Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00663.
XX
PR 04-FEB-2000; 2000US-0180312.
XX
PR 26-MAY-2000; 2000US-0207456.
XX
PR 30-JUN-2000; 2000US-0608408.
XX
PR 03-AUG-2000; 2000US-0632366.
XX
PR 21-SEP-2000; 2000US-0234687.
XX
PR 27-SEP-2000; 2000US-0236359.
XX
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488897/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human placenta -
XX
XX Claim 25; SEQ ID NO 6751; 654pp; English.
XX
XX The present invention relates to single exon nucleic acid probes (SENP).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders.
XX
SQ Sequence 475 BP; 134 A; 87 C; 86 G; 168 T; 0 other;
Query Match 11.7%; Score 36.4; DB 22; Length 475;

Best Local Similarity 52.7%; Pred. No. 1.3;
Matches 79; Conservative 0; Mismatches 71; Indels 0; Gaps 0;
OY 101 ATGTGTGATTCATATTCACCTAGTCGAATAATTGAATACATACTACCAATATACCATATATATCT 160
DB 53 ATGTACATTTTATATATATTTTAAATATTTATGTATACACACTACATACATATATCTAGAT 112
OY 161 AGACTGTATGTGTTCTTATACATATAGTATGACCTGAACTCCATTCAGTCAAGAAAAA 220
DB 113 ATACAGTAAATGATTATGCAATCTATTTTATATCTAGATCCAGAAATGTGGTAAACT 172
OY 221 TGGAGAAATAGCTATTTGTATCCATATGG 250
DB 173 GAGCACTATAATGAGAGTTTGTATCTGG 202

RESULT 15

ABS06984
ID ABS06984 standard; DNA; 475 BP.
AC ABS06984;
XX
DT 19-AUG-2002 (first entry)
XX
DE Human genome-derived single exon probe from lung SEQ ID No 6975.
XX
KW Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesia; pulmonary hypertension;
KW hyaline membrane disease.
XX
OS Homo sapiens.
PN
XX WO200186003-A2.
XX
PD 15-NOV-2001.
XX
XX 30-JAN-2001; 2001WO-US00665.
XX
PF 04-FEB-2000; 2000US-180312P.
PR 26-MAY-2000; 2000US-207456P.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-234687P.
PR 27-SEP-2000; 2000US-236359P.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.

Penn SG, Hanzel DK, Chen W, Rank DR;

WPI; 2002-114183/15.

Spatially-addressable set of single exon nucleic acid probes, used to measure gene expression in human lung samples -

Claim 1; SEQ ID No 6975; 634pp; English.

The invention relates to a spatially-addressable set of single exon nucleic acid probes for measuring gene expression in a sample derived from human lung comprising single exon nucleic acid probes having one of 12614 nucleic acid sequences mentioned in the specification, or their complements or the 12387 open reading frames derived from the 12614 probes. Also included are a microarray comprising the novel set of probes; the novel set of probes which hybridize at high stringency to a nucleic acid expressed in the human lung; measuring gene expression in a sample derived from human lung, comprising (a) contacting the array with a collection of detectably labeled nucleic acids derived from human lung

CC mRNA, and (b) measuring the label detectably bound to each probe of
CC the array; identifying exons in a eukaryotic genome, comprising
CC (a) algorithmically predicting at least one exon from genomic sequences
CC of the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene
CC expression analysis, and for identifying exons in a gene, particularly
CC using human lung derived mRNA and for the study of lung diseases
CC such as asthma, lung cancer, chronic obstructive pulmonary disease
CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic
CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
CC and hyaline membrane disease. The present sequence is a single exon
CC probe of the invention.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX

Sequence 475 BP; 134 A; 87 C; 86 G; 168 T; 0 other;

Query Match 11.7%; Score 36.4; DB 24; Length 475;
Best Local Similarity 52.7%; Pred. No. 1.3;
Matches 79; Conservative 0; Mismatches 71; Indels 0; Gaps 0;
OY 101 ATGTGTGATTCATATTCACCTAGTCGAATAATTGAATACATACTACCAATATATATCT 160
DB 53 ATGTACATTTTATATATATTTTAAATATTTATGTATACACACTACATATATCTAGAT 112
OY 161 AGACTGTATGTGTTCTTATACATATAGTATGACCTGAACTCCATTCAGTCAAGAAAAA 220
DB 113 ATACAGTAAATGATTATGCAATCTATTTTAAATATCTAGATCCAGAAATGTGGTAAACT 172
OY 221 TGGAGAAATAGCTATTTGTATCCATATGG 250
DB 173 GAGCACTATAATGAGAGTTTGTATCTGG 202

Search completed: July 7, 2003, 18:31:20

Job time : 221 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 7, 2003, 18:28:01 ; Search time 136 Seconds
(without alignments)
3569.620 Million cell updates/sec

Title: US-09-820-425B-12
Perfect score: 311
Sequence: 1 actcagctctgtgtgcaag.....gccattgcaggtaaagctt 311

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1085931 seqs, 780495707 residues
Total number of hits satisfying chosen parameters: 2171862

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	311	100.0	311	10	US-09-820-425B-12
2	36.4	11.7	475	10	US-09-864-761-11201
3	36.4	11.7	598	10	US-09-864-761-14718
4	36.4	11.7	2000	9	US-09-938-842A-3039
5	35	11.3	9310	7	US-08-781-986A-316
6	34.6	11.1	172637	10	US-09-805-458A-3
7	34.4	11.1	3425	9	US-10-073-961-315
8	34.4	11.1	3425	10	US-09-764-887-315
9	34.4	11.1	6945	9	US-10-073-961-313
10	34.4	11.1	6945	9	US-10-073-961-314
11	34.4	11.1	6945	10	US-09-764-887-313
12	34.4	11.1	6945	10	US-09-764-887-314
13	34.2	11.0	389	9	US-10-092-154-207
14	34.2	11.0	389	10	US-09-764-847-207
15	34.2	11.0	4579	9	US-10-273-680-6
16	34	10.9	143068	10	US-09-967-768A-316
17	33.8	10.9	200	10	US-09-867-701-8772
18	33.6	10.8	2000	9	US-09-938-842A-4431
19	33.6	10.8	53332	9	US-10-224-562-3

C 20	33.6	10.8	53332	10	US-09-801-861-3	Sequence 3, Appli
C 21	33.4	10.7	119040	9	US-09-911-077A-19	Sequence 19, Appl
C 22	33.2	10.7	640681	10	US-09-790-988-1	Sequence 1, Appli
C 23	33	10.6	436	9	US-10-102-524-1006	Sequence 1006, Ap
C 24	33	10.6	520	9	US-10-102-524-696	Sequence 696, App
C 25	33	10.6	535	9	US-10-102-524-501	Sequence 501, App
C 26	33	10.6	566	9	US-10-102-524-1642	Sequence 1642, Ap
C 27	33	10.6	1484	9	US-09-938-842A-2765	Sequence 2765, Ap
C 28	33	10.6	2429	10	US-09-838-044-3	Sequence 3, Appli
C 29	32.8	10.5	186957	9	US-10-185-770-3	Sequence 3, Appli
C 30	32.6	10.5	434	9	US-10-040-739-1284	Sequence 1284, Ap
C 31	32.6	10.5	596	10	US-09-864-761-12836	Sequence 12836, A
C 32	32.6	10.5	17672	9	US-10-144-649A-747	Sequence 747, App
C 33	32.6	10.5	31728	9	US-10-114-170-64	Sequence 64, Appl
C 34	32.6	10.5	161280	9	US-10-144-649A-746	Sequence 746, App
C 35	32.4	10.4	1079	9	US-09-443-218-1	Sequence 1, Appli
C 36	32.4	10.4	1452	10	US-09-969-708-275	Sequence 275, App
C 37	32.4	10.4	1849	9	US-09-893-519A-145	Sequence 145, App
C 38	32.4	10.4	9183	9	US-09-764-891-7135	Sequence 7135, Ap
C 39	32.4	10.4	127197	9	US-09-754-853A-1	Sequence 1, Appli
C 40	32.2	10.4	579	10	US-09-864-761-9873	Sequence 9873, Ap
C 41	32.2	10.4	975	9	US-10-081-051-20	Sequence 20, Appl
C 42	32.2	10.4	6190	9	US-10-081-051-19	Sequence 19, Appl
C 43	32.2	10.4	1691139	9	US-10-067-514-1	Sequence 1, Appli
C 44	32	10.3	346	9	US-10-046-935-241	Sequence 241, App
C 45	32	10.3	346	9	US-09-878-178-241	Sequence 241, App

ALIGNMENTS

RESULT 1
US-09-820-425B-12
; Sequence 12, Application US/09820425B
; Patent No. US20020119156A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Sei-Yu
; APPLICANT: Hu, Ping
; APPLICANT: Recipon, Herve
; APPLICANT: Macina, Roberto
; TITLE OF INVENTION: Compositions and Methods of Diagnosing, Monitoring,
; FILE REFERENCE: Staging, Imaging and Treating Lung Cancer
; FILE REFERENCE: DEX-0203
; CURRENT APPLICATION NUMBER: US/09/820,425B
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/192,921
; PRIOR FILING DATE: 2000-03-29
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 311
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-820-425B-12

Query Match	100.0%	Score 311;	DB 10;	Length 311;
Best Local Similarity	100.0%;	Pred. No. 2.1e-76;		
Matches 311;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ACTCAGCTCTGTGTGCAAGGAGATGCTGTGGAATGTCACAGCATCGTATAGCAAGAGC	60	
Db	1	ACTCAGCTCTGTGTGCAAGGAGATGCTGTGGAATGTCACAGCATCGTATAGCAAGAGC	60	
QY	61	ATATTGGCAACAGCTTGGATGCCAGCAGAGAGGCCCAAAATGTGTGATTTCATATTTCACT	120	
Db	61	ATATTGGCAACAGCTTGGATGCCAGCAGAGAGGCCCAAAATGTGTGATTTCATATTTCACT	120	
QY	121	AGTCGAATAATTAATCACTACAAATATACACCATATATACAGCTGTATGTTGTTCTTA	180	
Db	121	AGTCGAATAATTAATCACTACAAATATACACCATATATACAGCTGTATGTTGTTCTTA	180	
QY	181	TACTATAGTGATTCCTTGAACCTCCACTCAGTGAAGAAAATGGAAGAAATAGCTATTTGT	240	

Db 181 TACTATAGTATTGACTTGACCTCCATTTCAGTCAAAATGGAAGAAATGAGTATTGT 240
QY 241 ATCCATATGGATACAAAAAGCAGGGTACAAAAGAAATACATCATCTTGCCATTGC 300
|||||
Db 241 ATCCATATGGATACAAAAAGCAGGGTACAAAAGAAATACATCATCTTGCCATTGC 300
QY 301 AGGTAAGCTT 311
|||||
Db 301 AGGTAAGCTT 311

RESULT 2

US-09-864-761-11201
; Sequence 11201, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeonica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 11201
; LENGTH: 475
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005411.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1

; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.99
US-09-864-761-11201

Query Match 11.7%; Score 36.4; DB 10; Length 475;
Best Local Similarity 52.7%; Pred. No. 2.2;
Matches 79; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 101 ATGTGTGATTTCATATTCACCTAGTCGAAATAATTCGAATACATAATACACCATATATACT 160
|||||
Db 53 ATGTACTTTTATATTAATTTTAAATATATGATATACACTACTATACATATATAGAT 112
QY 161 AGACTGTATGTGTGTTCTATATACTATATAGTGAATGACTTGAATCCATTCAGTCAAAAAA 220
|||||
Db 113 ATACAGTAAATGATTATGCTTCTATTTAAATATCTAGATCCAGAATGTGGGTAAACT 172
QY 221 TGGAGAANTAGCTATTTTGTATCCATATGG 250
|||||
Db 173 GAGCACTAATAATGAGAGATTTAGTAATCTGG 202

RESULT 3

US-09-864-761-14718
; Sequence 14718, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aeonica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1

; SEQ ID NO 14718

; LENGTH: 598

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: MAP TO AC015842.4

; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.55

; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.7

; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.69

; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.51

US-09-864-761-14718

Query Match

Best Local Similarity 11.7%; Score 36.4; DB 10; Length 598;

Matches 79; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 101 ATGTGTTGATTCATATTCAGTTCGAAATTAATGATACATACATACATATATACT 160

Db 45 ATGACTTTTATATTAATTTTATATTTATATTTATATGATACATACATATATCTAGAT 104

QY 161 AGACTGTATGTGTTCTTCTATACATAGTGTGACTTGAACCTCCATTCAGTGAAGAAA 220

Db 105 ATACAGTAAGATGATATCATCTCTATTTTATATCTAGATCCAGATGGTAAACT 164

QY 221 TGGAGAATAGCTATTTGTATCCATATGG 250

Db 165 GAGCACTATAATGAGAGTTAGTAATCTGG 194

RESULT 4

US-09-938-842A-3039/c

; Sequence 3039, Application US/09938842A

; Patent No. US20020160378A1

; GENERAL INFORMATION:

; APPLICANT: Harper, Jeff

; APPLICANT: Kreps, Joel

; APPLICANT: Wang, Xun

; APPLICANT: Zhu, Tong

; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING

; FILE REFERENCE: SCRIPI300-3

; CURRENT APPLICATION NUMBER: US/09/938,842A

; CURRENT FILING DATE: 2001-08-24

; PRIOR APPLICATION NUMBER: US 60/227,866

; PRIOR FILING DATE: 2000-08-24

; PRIOR APPLICATION NUMBER: US 60/264,647

; PRIOR FILING DATE: 2001-01-16

; PRIOR APPLICATION NUMBER: US 60/300,111

; PRIOR FILING DATE: 2001-06-22

; NUMBER OF SEQ ID NOS: 5379

; SEQ ID NO 3039

; LENGTH: 2000

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

US-09-938-842A-3039

Query Match

Best Local Similarity 11.7%; Score 36.4; DB 9; Length 2000;

Matches 82; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 101 ATGTGTTGATTCATATTCAGTTCGAAATTAATGATACATACATACATATATACT 160

Db 805 ATATATTTTATGTTTATATTTTCTAAAGTTCATTTTATATTTTACTATATTCAAA 746

QY 161 AGACTGTATGTGTTCTTCTATACATAGTGTGACTTGAACCTCCATTCAGTGAAGAAA 220

Db 745 ATCTTTTGTGTTTCTTCTATACATAATATTTTATTTAGTGTGTAATAGTCTTCTAAA 686

QY 221 TGGAGAATAGCTATTTGTATCCATATGGATACAAA 258

Db 685 AATAATAATTATACATAGTATTTTAAAAACATAAA 648

RESULT 5

US-08-781-986A-316

; Sequence 316, Application US/08781986A

; Publication No. US20030054436A1

; GENERAL INFORMATION:

; APPLICANT: Charles Kunsch

; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences

; NUMBER OF SEQUENCES: 5255

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Human Genome Sciences, Inc.

; STREET: 9410 Key West Avenue

; CITY: Rockville

; STATE: Maryland

; COUNTRY: USA

; ZIP: 20850

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

; COMPUTER: HP Vectra 486/33

; OPERATING SYSTEM: MSDOS version 6.2

; SOFTWARE: ASCII Text

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/781,986A

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Benson, Bob

; REGISTRATION NUMBER: 30,446

; REFERENCE/DOCKET NUMBER: PB248PP

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (301) 309-8504

; TELEFAX: (301) 309-8512

; INFORMATION FOR SEQ ID NO: 316:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 9310 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

US-08-781-986A-316

Query Match

Best Local Similarity 11.3%; Score 35; DB 7; Length 9310;

Matches 83; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 143 ATATACACCATATATAGTACTGTATGTTTCTTACTATATAGTACTGACTTGAAC 202

Db 5558 ATTTATATGATTTTAAAGCGACGTTGAAGCGGCACAAAAATTTATGACTTATTAAAC 5617

QY 203 TCCATTTCAGTGAAGAAATGGAAGATTTAGCTATTGTTATCATATGGATACAAAAAG 262

Db 5618 CTATTTTAGAGAAAAAGATAAAATTAAGTGATGATATCCAAATGAATTCGATAAG 5677

QY 263 CAGGTAAACAAAAGAAATCTACATCATCTTGGCATTTCGAGGTA 305

Db 5678 TGAATCAATTATTCGATAAATAAAGATAACACGCGCGTTA 5720

RESULT 6

US-09-805-458A-3

; Sequence 3, Application US/09805458A

; Patent No. US20020042100A1

; GENERAL INFORMATION:

; APPLICANT: YAN, Chunhua et al

; TITLE OF INVENTION: ISOLATED HUMAN ION CHANNEL PROTEINS,

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN ION CHANNEL PROTEINS,

; TITLE OF INVENTION: AND USES THEREOF

; FILE REFERENCE: CL000722

; CURRENT APPLICATION NUMBER: US/09/805,458A

; CURRENT FILING DATE: 2001-03-14

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: FastSeq for Windows Version 4.0


```

; PRIOR APPLICATION NUMBER: 60/232,401
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/241,808
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,826
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,786
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,221
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,475
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/231,243
; PRIOR FILING DATE: 2000-09-08

Query Match          11.1%; Score 34.4; DB 9; Length 3425;
Best Local Similarity 50.6%; Pred. No. 21;
Matches 83; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 134 AATACTACAATATACACCATATATACACTAGACTGTATGTGTCTATATACTATAGTGATT 193
Db 1068 AGTGCCGCATTAACCCCAATACAGTTGGCTCTCTATATCTGTGGATTATGCACCTGGT 1009

QY 194 GACTTGAACCTCAATTCAGTGAAGAAAATGGAAGAATTAGCTATTTGTATCCATATGGGAT 253
Db 1008 AGATTCAACCAATTTTGGATAGAAAATATCTACGAAAAACAATTTAAAAAACTACAACCT 949

QY 254 ACAAAAAGCAGGTTACAAAAGAATCTACATCATCTTGGCAAT 297
Db 948 AAAAATATATATAGTAAAATACTATTTACATAGCATTTACATT 905

RESULT 8
US-09-764-887-315/C
; Sequence 315, Application US/09764887
; Patent No. US20020042096A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PAL13
; CURRENT APPLICATION NUMBER: US/09/764,887
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 658
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 315
; LENGTH: 3425
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-887-315

Query Match          11.1%; Score 34.4; DB 10; Length 3425;
Best Local Similarity 50.6%; Pred. No. 21;
Matches 83; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 134 AATACTACAATATACACCATATATACACTAGACTGTATGTGTCTATATACTATAGTGATT 193
Db 1068 AGTGCCGCATTAACCCCAATACAGTTGGCTCTCTATATCTGTGGATTATGCACCTGGT 1009

QY 194 GACTTGAACCTCAATTCAGTGAAGAAAATGGAAGAATTAGCTATTTGTATCCATATGGGAT 253
Db 1008 AGATTCAACCAATTTTGGATAGAAAATATCTACGAAAAACAATTTAAAAAACTACAACCT 949

QY 254 ACAAAAAGCAGGTTACAAAAGAATCTACATCATCTTGGCAAT 297
Db 948 AAAAATATATATAGTAAAATACTATTTACATAGCATTTACATT 905

RESULT 9
US-10-073-961-313/c
; Sequence 313, Application US/10073961
; Patent No. US20030077602A1
; GENERAL INFORMATION:

```

APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: P4113C1
CURRENT APPLICATION NUMBER: US/10/073,961
CURRENT FILING DATE: 2002-03-05
PRIOR APPLICATION NUMBER: 09/764,887
PRIOR FILING DATE: 2001-01-17
PRIOR APPLICATION NUMBER: 60/179,065
PRIOR FILING DATE: 2000-01-31
PRIOR APPLICATION NUMBER: 60/180,628
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: 60/214,886
PRIOR FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: 60/217,487
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225,758
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/220,963
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: 60/217,496
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225,447
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/218,290
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: 60/225,757
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/226,868
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/216,647
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: 60/225,267
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/216,880
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: 60/225,270
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/251,869
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/235,834
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/234,274
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: 60/234,223
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: 60/228,924
PRIOR FILING DATE: 2000-08-30
PRIOR APPLICATION NUMBER: 60/224,518
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/236,369
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/224,519
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/220,964
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: 60/241,809
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/249,299
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/236,327
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/241,785
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/244,617
PRIOR FILING DATE: 2000-11-01
PRIOR APPLICATION NUMBER: 60/225,268
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/236,368
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/251,856
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/251,868
PRIOR FILING DATE: 2000-12-08

PRIOR APPLICATION NUMBER: 60/229,344
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/234,997
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: 60/229,343
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,345
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,287
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,513
PRIOR FILING DATE: 2000-09-05
PRIOR APPLICATION NUMBER: 60/231,413
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/229,509
PRIOR FILING DATE: 2000-09-05
PRIOR APPLICATION NUMBER: 60/236,367
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/237,039
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237,038
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/236,370
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/236,802
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237,037
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237,040
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/240,960
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/239,935
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/239,937
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/241,787
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/246,474
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/246,532
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/249,216
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,210
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/226,681
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/225,759
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/225,213
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/227,182
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/225,214
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/235,836
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/230,438
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/215,135
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: 60/225,266
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/249,218
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,208
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,213
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,212
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,207

1	1	PRIOR FILING DATE: 2000-12-08	1
2	2	PRIOR APPLICATION NUMBER: 60/229,344	2
3	3	PRIOR FILING DATE: 2000-09-01	3
4	4	PRIOR APPLICATION NUMBER: 60/234,997	4
5	5	PRIOR FILING DATE: 2000-09-25	5
6	6	PRIOR APPLICATION NUMBER: 60/229,343	6
7	7	PRIOR FILING DATE: 2000-09-01	7
8	8	PRIOR APPLICATION NUMBER: 60/229,345	8
9	9	PRIOR FILING DATE: 2000-09-01	9
10	10	PRIOR APPLICATION NUMBER: 60/229,287	10
11	11	PRIOR FILING DATE: 2000-09-01	11
12	12	PRIOR APPLICATION NUMBER: 60/229,513	12
13	13	PRIOR FILING DATE: 2000-09-05	13
14	14	PRIOR APPLICATION NUMBER: 60/231,413	14
15	15	PRIOR FILING DATE: 2000-09-08	15
16	16	PRIOR APPLICATION NUMBER: 60/229,509	16
17	17	PRIOR FILING DATE: 2000-09-05	17
18	18	PRIOR APPLICATION NUMBER: 60/236,367	18
19	19	PRIOR FILING DATE: 2000-09-29	19
20	20	PRIOR APPLICATION NUMBER: 60/237,039	20
21	21	PRIOR FILING DATE: 2000-10-02	21
22	22	PRIOR APPLICATION NUMBER: 60/237,038	22
23	23	PRIOR FILING DATE: 2000-10-02	23
24	24	PRIOR APPLICATION NUMBER: 60/236,370	24
25	25	PRIOR FILING DATE: 2000-09-29	25
26	26	PRIOR APPLICATION NUMBER: 60/236,802	26
27	27	PRIOR FILING DATE: 2000-10-02	27
28	28	PRIOR APPLICATION NUMBER: 60/237,037	28
29	29	PRIOR FILING DATE: 2000-10-02	29
30	30	PRIOR APPLICATION NUMBER: 60/237,040	30
31	31	PRIOR FILING DATE: 2000-10-02	31
32	32	PRIOR APPLICATION NUMBER: 60/240,960	32
33	33	PRIOR FILING DATE: 2000-10-20	33
34	34	PRIOR APPLICATION NUMBER: 60/239,935	34
35	35	PRIOR FILING DATE: 2000-10-13	35
36	36	PRIOR APPLICATION NUMBER: 60/239,937	36
37	37	PRIOR FILING DATE: 2000-10-13	37
38	38	PRIOR APPLICATION NUMBER: 60/241,787	38
39	39	PRIOR FILING DATE: 2000-10-20	39
40	40	PRIOR APPLICATION NUMBER: 60/246,474	40
41	41	PRIOR FILING DATE: 2000-11-08	41
42	42	PRIOR APPLICATION NUMBER: 60/246,532	42
43	43	PRIOR FILING DATE: 2000-11-08	43
44	44	PRIOR APPLICATION NUMBER: 60/249,216	44
45	45	PRIOR FILING DATE: 2000-11-17	45
46	46	PRIOR APPLICATION NUMBER: 60/249,210	46
47	47	PRIOR FILING DATE: 2000-11-17	47
48	48	PRIOR APPLICATION NUMBER: 60/226,681	48
49	49	PRIOR FILING DATE: 2000-08-22	49
50	50	PRIOR APPLICATION NUMBER: 60/225,759	50
51	51	PRIOR FILING DATE: 2000-08-14	51
52	52	PRIOR APPLICATION NUMBER: 60/225,213	52
53	53	PRIOR FILING DATE: 2000-08-14	53
54	54	PRIOR APPLICATION NUMBER: 60/227,182	54
55	55	PRIOR FILING DATE: 2000-08-22	55
56	56	PRIOR APPLICATION NUMBER: 60/225,214	56
57	57	PRIOR FILING DATE: 2000-08-14	57
58	58	PRIOR APPLICATION NUMBER: 60/235,836	58
59	59	PRIOR FILING DATE: 2000-09-27	59
60	60	PRIOR APPLICATION NUMBER: 60/230,438	60
61	61	PRIOR FILING DATE: 2000-09-06	61
62	62	PRIOR APPLICATION NUMBER: 60/215,135	62
63	63	PRIOR FILING DATE: 2000-06-30	63
64	64	PRIOR APPLICATION NUMBER: 60/225,266	64
65	65	PRIOR FILING DATE: 2000-08-14	65
66	66	PRIOR APPLICATION NUMBER: 60/249,218	66
67	67	PRIOR FILING DATE: 2000-11-17	67
68	68	PRIOR APPLICATION NUMBER: 60/249,208	68
69	69	PRIOR FILING DATE: 2000-11-17	69
70	70	PRIOR APPLICATION NUMBER: 60/249,213	70
71	71	PRIOR FILING DATE: 2000-11-17	71
72	72	PRIOR APPLICATION NUMBER: 60/249,212	72
73	73	PRIOR FILING DATE: 2000-11-17	73

1	PRIOR APPLICATION NUMBER: 60/249,207
2	PRIOR FILING DATE: 2000-11-17
3	PRIOR APPLICATION NUMBER: 60/249,245
4	PRIOR FILING DATE: 2000-11-17
5	PRIOR APPLICATION NUMBER: 60/249,244
6	PRIOR FILING DATE: 2000-11-17
7	PRIOR APPLICATION NUMBER: 60/249,217
8	PRIOR FILING DATE: 2000-11-17
9	PRIOR APPLICATION NUMBER: 60/249,211
10	PRIOR FILING DATE: 2000-11-17
11	PRIOR APPLICATION NUMBER: 60/249,215
12	PRIOR FILING DATE: 2000-11-17
13	PRIOR APPLICATION NUMBER: 60/249,264
14	PRIOR FILING DATE: 2000-11-17
15	PRIOR APPLICATION NUMBER: 60/249,214
16	PRIOR FILING DATE: 2000-11-17
17	PRIOR APPLICATION NUMBER: 60/249,297
18	PRIOR FILING DATE: 2000-11-17
19	PRIOR APPLICATION NUMBER: 60/232,400
20	PRIOR FILING DATE: 2000-09-14
21	PRIOR APPLICATION NUMBER: 60/231,242
22	PRIOR FILING DATE: 2000-09-08
23	PRIOR APPLICATION NUMBER: 60/232,081
24	PRIOR FILING DATE: 2000-09-08
25	PRIOR APPLICATION NUMBER: 60/232,080
26	PRIOR FILING DATE: 2000-09-08
27	PRIOR APPLICATION NUMBER: 60/231,414
28	PRIOR FILING DATE: 2000-09-08
29	PRIOR APPLICATION NUMBER: 60/231,244
30	PRIOR FILING DATE: 2000-09-08
31	PRIOR APPLICATION NUMBER: 60/233,064
32	PRIOR FILING DATE: 2000-09-14
33	PRIOR APPLICATION NUMBER: 60/233,063
34	PRIOR FILING DATE: 2000-09-14
35	PRIOR APPLICATION NUMBER: 60/232,397
36	PRIOR FILING DATE: 2000-09-14
37	PRIOR APPLICATION NUMBER: 60/232,399
38	PRIOR FILING DATE: 2000-09-14
39	PRIOR APPLICATION NUMBER: 60/232,401
40	PRIOR FILING DATE: 2000-09-14
41	PRIOR APPLICATION NUMBER: 60/241,808
42	PRIOR FILING DATE: 2000-10-20
43	PRIOR APPLICATION NUMBER: 60/241,826
44	PRIOR FILING DATE: 2000-10-20
45	PRIOR APPLICATION NUMBER: 60/241,786
46	PRIOR FILING DATE: 2000-10-20
47	PRIOR APPLICATION NUMBER: 60/241,221
48	PRIOR FILING DATE: 2000-10-20
49	PRIOR APPLICATION NUMBER: 60/246,475
50	PRIOR FILING DATE: 2000-11-08
51	PRIOR APPLICATION NUMBER: 60/231,243
52	PRIOR FILING DATE: 2000-09-08

Query Match	Score 34.4;	DB 9;	Length 6945;
11.18;			

Best Local Similarity 50.6%; Pred. No. 29;

Best Local Similarity 50.00; Recd. No: 25;
Matches 83; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

134 22TCTCACAATATACCCATATATACCTAGCATCTCTATCTCTCTCTCTCTATAGCTGATT 193

— — — — —

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

DB 4528 AGATTCAACCAATTTGGATAGAAAATATCTACGAAAAACAATTTAAATAACTACAACTT 4469

QY 254 ACAAAAAGCAGGGTAACAAAGAATCTACATCATCTTGCCATT 297

RESULT 11

RESULT II
NS-09-76A-887-313/c

US=09-/64-88/-313/c
: Sequence 313 Application US/09764887

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 7, 2003, 18:22:48 ; Search time 1424 Seconds
(without alignments)
3537.076 Million cell updates/sec

Title: US-09-820-425B-12

Perfect score: 311

Sequence: 1 actccagctctgtgtgcaag.....gccattgcaggtaaagctt 311

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:**

2: em_esthum:**

3: em_estin:**

4: em_estmu:**

5: em_estor:**

6: em_estpl:**

7: em_estro:**

8: em_htc:**

9: gb_est1:**

10: gb_est2:**

11: gb_htc:**

12: gb_est3:**

13: gb_est4:**

14: gb_est5:**

15: em_estfun:**

16: em_estom:**

17: gb_gss:**

18: em_gss_hum:**

19: em_gss_inv:**

20: em_gss_pln:**

21: em_gss_vrt:**

22: em_gss_fun:**

23: em_gss_mam:**

24: em_gss_mus:**

25: em_gss_othr:**

26: em_gss_pro:**

27: em_gss_rod:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	43.8	14.1	881	17	AZ531774 ENTBJ30TR
2	43.8	14.1	923	17	AZ529975 ENTDK18TF
3	41.6	13.4	837	17	BH152288 ENTQNS7TF
4	41.6	13.4	872	17	BH137298 ENTBN94TR
5	41.4	13.3	443	12	BF857983 RC5-FT019
6	40.8	13.1	512	17	AQ673535 HS_5491_A

7	40.2	12.9	350	14	R23049	R23049 yh31a01.r1
8	40.2	12.9	457	10	BE148432	BE148432 MRO-HF024
9	39.2	12.6	564	17	AQ682268	AQ682268 HS_5505_B
10	39	12.5	383	10	BE111015	BE111015 UI-R-BJ1-
11	39	12.5	415	9	AA964465	AA964465 UI-R-EI-g
12	39	12.5	476	10	BE103314	BE103314 UI-R-BX0-
13	39	12.5	516	12	BF397229	BF397229 UI-R-BS2-
14	39	12.5	720	14	BQ192595	BQ192595 UI-R-DR1-
15	38.8	12.5	698	17	AQ367250	AQ367250 toxb0002E
16	38.8	12.5	713	17	AQ367260	AQ367260 toxb0002E
17	38.6	12.4	287	17	AQ029908	AQ029908 RPC111-37
18	38.6	12.4	608	17	AQ281039	AQ281039 RPC111-88
19	38.6	12.4	784	17	AZ286163	AZ286163 RPC11-23-1
20	38.6	12.4	877	17	CNS012VL	AL102075 Drosophila
21	38.6	12.4	886	17	CNS01NM2	AL152755 Anopheles
22	38.4	12.3	770	17	BH401807	BH401807 AG-ND-102
23	38.4	12.3	916	17	CNS041QP	AL292570 Tetraodon
24	38.2	12.3	240	14	C56318	C56318 C56318 YuJ1
25	38.2	12.3	554	12	BG839684	BG839684 Gm01_05d0
26	38.2	12.3	878	17	CNS0187R	AL108993 Drosophila
27	38	12.2	472	12	BG447078	BG447078 GA_EB004
28	38	12.2	963	17	BH134011	BH134011 ENTOH02TF
29	37.6	12.1	529	17	AZ965766	AZ965766 2M0236D04
30	37.6	12.1	852	17	CNS021J2	AL178951 Tetraodon
31	37.4	12.0	529	17	B89562	B89562 RPC111-20G1
32	37.4	12.0	703	17	AG169709	AG169709 Pan trogl
33	37.4	12.0	970	17	CNS0182E	AL108800 Drosophila
34	37.4	12.0	1084	17	CNS071NH	AL425139 clone_BAO
35	37.4	12.0	1101	17	CNS00LVP	AL078809 Drosophila
36	37.2	12.0	419	13	BM275039	BM275039 PEST00a7
37	37.2	12.0	527	13	BM169030	BM169030 EST571553
38	37.2	12.0	1101	17	CNS00396	AL063911 Drosophila
39	37	11.9	372	13	BG993266	BG993266 MR3-HT110
40	37	11.9	459	17	AZ867858	AZ867858 2M0178C22
41	37	11.9	480	13	BI513391	BI513391 BB160012B
42	37	11.9	914	17	AZ538928	AZ538928 ENTFR23TR
43	37	11.9	920	17	BH164585	BH164585 ENTFR27TR
44	37	11.9	1034	17	CNS0166R	AL106365 Drosophila
45	36.8	11.8	500	9	AU087152	AU087152 AU087152

ALIGNMENTS

RESULT 1
AZ531774
LOCUS ENTBJ30TR Entamoeba histolytica Sheared DNA linear 881 bp GSS 03-NOV-2000
DEFINITION genomic, DNA sequence.
ACCESSION AZ531774
VERSION AZ531774.1 GI:11086784
KEYWORDS GSS.
SOURCE Entamoeba histolytica.
ORGANISM Entamoeba histolytica
REFERENCE Eukaryota; Entamoebidae; Entamoeba.
AUTHORS Loftus, B., Van Aken, S. and Fraser, C.
TITLE Determination of clone end sequences from Entamoeba histolytica
JOURNAL HMI:IMSS sheared DNA library
COMMENT Unpublished (2000)
Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 3543
Email: b.loftus@tigr.org
Clones are derived from the Entamoeba histolytica HMI:IMSS sheared
DNA library
Seq primer: M13-Reverse
Class: shotgun
High quality sequence start: 14
High quality sequence stop: 786.


```

/sex="Female"
/dev_stage="placenta obtained at birth (full term)"
/lab_host="DH10B (ampicillin resistant)"
/notes="Organ: placenta; Vector: pT7T3D (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5',
AACTGGAAGAAATCGCGCCGAGGAATTTTTTTTTTTTTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo. "
```

```

BASE COUNT      114 a      56 c      45 g      130 t      5 others
ORIGIN

Query Match      12.9%; Score 40.2; DB 14; Length 350;
Best Local Similarity 51.7%; Pred. No. 2.8;
Matches 90; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 134 AATACATACATATACACCATATATATAGACGTATGTGTTCTATATAGTATGATT 193
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 107 ATTTTAAAAACAAGCATTTGGATTGTTTCACTGTAAGTCTTGTACATCTTAGGATTGAT 166
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 194 GACTTCACCTCATTTCAGTGAAGAAATGGAAGATTACCTATTGTTATCCATATGGGAT 253
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 167 GCGTTAAACTTCCAAGACTTGAAGAAATAGTGAAGAAATTTTACTAGGAAATATTTTGC 226
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 254 ACAAAAGACGAGTAAACAAAGAAATCTCATCTTGCCTATTTCAGGTAA 307
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 227 ATATATATTTAGGGCAATATATATATATATATATATATATATATATATATATATAT 280
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```

RESULT 8
LOCUS      BE148432
DEFINITION MR0-HT0241-120100-005-e04 HT0241 Homo sapiens cDNA, mRNA sequence.
ACCESSION  BE148432
VERSION     BE148432.1 GI:8611156
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
```

```

REFERENCE   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS     Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
            1 (bases 1 to 457)
            Dias Neto,E., Garcia Correa,R., Verjovsky-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
```

```

TITLE       Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
JOURNAL      Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE      20202663
COMMENT      Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
```

```

This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2=MR0-HT0241-120
100-005-e04&st3=2000-01-12&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 33
High quality sequence stop: 216.
```

```

FEATURES
            Location/Qualifiers
            1..457
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
```

```

/clone_lib="HT0241"
/dev_stage="Adult"
/notes="Organ: head neck; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT      142 a      102 c      80 g      133 t
ORIGIN
```

```

Query Match      12.9%; Score 40.2; DB 10; Length 457;
Best Local Similarity 53.5%; Pred. No. 2.9;
Matches 84; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 30 TGGATGTCACAGCATCGTATAGCAAGAGCATATTTGGCAACAGCTTGGATGCCAGCAG 89
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 214 TGGGTATATATCCCATTTATATAGCTTTCAAAATATTTGGAGACACCATGTTTGGAAAGGTA 273
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 90 AAGGAGCCCCAAATGTGTGATTCATATTCATCTACTAGTGAATAATTGAATCTACATAATACA 149
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 274 AAACATATAGCATATGAGACACCTGTTTATCATAGTAGAGCAATTAACAACTAGCTAGTGA 333
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 150 CCATATATATAGCATGTATGTGTTTCTTATATAT 186
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 334 CCAAACTGCCCAACTCTATGTATATATAGTTTCTAT 370
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```

RESULT 9
LOCUS      AQ682268/c
DEFINITION HS_5505_B2_H04_T7A RPCI-11 Human Male BAC Library Homo sapiens
ACCESSION  AQ682268
VERSION     AQ682268.1 GI:5258251
KEYWORDS    GSS.
SOURCE      human.
ORGANISM    Homo sapiens
```

```

REFERENCE   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS     Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
            1 (bases 1 to 564)
            Mahairs,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
```

```

TITLE       Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
JOURNAL      Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
MEDLINE      99380589
COMMENT      Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
```

```

Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact pieter de Jong
(pieter@edjong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
or from Resear h Genetics (info@resgen.com). BAC end Web Server:
http://www.htsc.washington.edu
Plate: 1081 row: P column: 8
Seq primer: T7
```

```

Class: BAC ends
High quality sequence stop: 564.
            Location/Qualifiers
            1..564
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone_lib="RPCI-11 Human Male BAC Library"
            /sex="male"
```

```

FEATURES
            source
```


and the oligo-dT track served to verify it as a clone from the normalized embryo at 13 dpc library cDNA Library Preparation: M.B. Soares Lab Clone distribution: clones will be available through Research Genetics (www.resgen.com)
Seq primer: M13 Forward
POLYA=Yes.

FEATURES
source
Location/Qualifiers
1. .516
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone_lib="UI-R-BS2-bej-h-08-0-UI"
/clone_lib="UI-R-BS2"
/dev_stage="embryonic 13 dpc"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; The UI-R-BS2 library is a subtracted library derived from 13 dpc whole embryo tissue. For a detailed description of the library from which this clone was derived, please visit our web site at ratest.eng.uiowa.edu. The subtraction has been previously described in (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)
TAG_LIB=UI-R-BS2
TAG_TISSUE=embryo at 13 dpc
TAG_SEQ=AATCC"
BASE COUNT 173 a 99 c 81 g 163 t
ORIGIN

Query Match 12.5%; Score 39; DB 12; Length 516;
Best Local Similarity 56.7%; Pred. No. 6.3;
Matches 72; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
QY 153 TATATACTAGACTGATGTGTTCTTATCTATAGTATGACTGAACTCAATTCAGT 212
||||| || ||||||| ||| || ||||||| ||| |||
Db 282 TATATGCTGCTGATGTATGTATGACATTTTCACAGTGATTCATTTGTAACGAA 223
QY 213 GAAAAAATGGAAGATTAGCTATTGTATCCATATGGATACAAAAAGCAGGGTAACA 272
||| || ||||||| ||| ||| ||| ||| ||| ||| |||
Db 222 AAAGACAAAAAAGAACATCACCACTGTCTCTATTTATGAAGATGATAAGCAGCTTTATT 163
QY 273 AAAGAT 279
||| ||| |||
Db 162 AATTAT 156

RESULT 14
BQ192595/c
LOCUS
DEFINITION
QY 192595 720 bp mRNA linear EST 30-APR-2002
UI-R-DRI-ckz-h-22-0-UI.s1 UI-R-DRI Rattus norvegicus cDNA clone
QY 192595
VERSION
KEYWORDS
SOURCE
ORGANISM
Norway rat.
Rattus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 720)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene discovery
Genome Res. 6 (9), 791-806 (1996)
9704477
Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.wiwi.edu
The sequence contained an oligo-dT track that was present in the

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
COMMENT

oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to verify it as a clone from the normalized osteoblast library cDNA Library Preparation: M.B. Soares Lab Clone distribution: clones will be available through Research Genetics (www.resgen.com)
Seq primer: M13 Forward
POLYA=Yes.

FEATURES
source
Location/Qualifiers

1. .720
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone_lib="UI-R-DRI-ckz-h-22-0-UI"
/clone_lib="UI-R-DRI"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; The UI-R-DRI library is a normalized Rat Osteoblast library (nREO) constructed in pT37 vector according to the procedure described by Bonaldo, Lennon & Soares (Normalization and Subtraction: Two Approaches to Facilitate Gene Discovery. Genome Research 6: 791-806, 1996). The oligonucleotide used to prime first strand synthesis contained the sequence tag AAGATATCAA between the Not I cloning site and the 18 stretch. The Rat Osteoblast tissue was provided by Lian & Stein of the University of Massachusetts Medical School.

TAG_LIB=UI-R-DRI
TAG_TISSUE=Osteoblast
TAG_SEQ=AAGATATCAA"
BASE COUNT 231 a 148 c 121 g 220 t
ORIGIN

Query Match 12.5%; Score 39; DB 14; Length 720;
Best Local Similarity 56.7%; Pred. No. 6.8;
Matches 72; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
QY 153 TATATACTAGACTGATGTGTTCTTATCTATAGTATGACTGAACTCAATTCAGT 212
||||| || ||||||| ||| || ||||||| ||| |||
Db 276 TATATGCTGCTGATGTATGTATGACATTTTCACAGTGATTCATTTGTAACGAA 217
QY 213 GAAAAAATGGAAGATTAGCTATTGTATCCATATGGATACAAAAAGCAGGGTAACA 272
||| ||| ||||||| ||| ||| ||| ||| ||| ||| |||
Db 216 AAAGACAAAAAAGAACATCACCACTGTCTCTATTTATGAAGATGATAAGCAGCTTTATT 157
QY 273 AAAGAT 279
||| ||| |||
Db 156 AATTAT 150

RESULT 15
AQ367250/c
LOCUS
DEFINITION
QY 367250 698 bp DNA linear GSS 07-MAR-2000
tox0002E12r CUGI Tomato BAC Library Lycopersicon esculentum genomic clone tox0002E12r, DNA sequence.
QY 367250
VERSION
KEYWORDS
SOURCE
ORGANISM
tomato.
Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
1 (bases 1 to 698)
Budiman,M.A., Mao,L., Wood,T.C. and Wing,R.A.
A deep-coverage tomato BAC library and prospects toward development of an STC framework for genome sequencing
Genome Res. 10 (1), 129-136 (2000)
20113122

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 7, 2003, 18:26:00 ; Search time 65 Seconds
(without alignments)
1467.330 Million cell updates/sec

Title: US-09-820-425B-12
Perfect score: 311
Sequence: 1 actccagctctgtgtgcaag.....gccatttcgaggtaaaagctt 311

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	35	11.3	72604	US-09-268-992-7	Sequence 7, Appli
C 2	35	11.3	72604	US-09-657-474-7	Sequence 7, Appli
C 3	33	10.6	2429	US-09-386-493-3	Sequence 3, Appli
C 4	33	10.6	5889	US-09-402-929-3	Sequence 3, Appli
C 5	32.6	10.5	1046	US-08-361-467B-4	Sequence 4, Appli
C 6	32.6	10.5	1046	US-08-484-332C-4	Sequence 4, Appli
C 7	32.6	10.5	1701	US-08-484-993B-40	Sequence 40, Appl
C 8	32.6	10.5	1701	US-08-484-158B-40	Sequence 40, Appl
C 9	32.6	10.5	1701	US-08-484-596A-40	Sequence 40, Appl
C 10	32.6	10.5	1701	US-08-480-150A-40	Sequence 40, Appl
C 11	32.6	10.5	1701	US-08-458-731-40	Sequence 40, Appl
C 12	32.6	10.5	1701	US-08-149-223A-40	Sequence 40, Appl
C 13	32.6	10.5	31728	US-09-453-702B-64	Sequence 64, Appl
C 14	32.6	10.5	168575	US-09-426-290-1	Sequence 1, Appli
C 15	32.4	10.4	1042	US-08-765-381-3	Sequence 3, Appli
C 16	32.4	10.4	1079	US-08-765-381-1	Sequence 1, Appli
C 17	31.8	10.2	3006	US-09-449-285A-1	Sequence 1, Appli
C 18	31.6	10.2	1545	US-08-400-275-17	Sequence 17, Appl
C 19	31.6	10.2	2466	US-09-363-243-1	Sequence 1, Appli
C 20	31.6	10.2	4459	US-09-363-243-2	Sequence 2, Appli
C 21	31.2	10.0	2943	US-09-503-922-2	Sequence 2, Appli
C 22	31	10.0	2755	US-08-749-522-2	Sequence 2, Appli
C 23	30.8	9.9	1365	US-08-870-827-4	Sequence 4, Appli
C 24	30.8	9.9	1365	US-09-317-179-4	Sequence 4, Appli
C 25	30.8	9.9	2408	US-08-870-827-5	Sequence 5, Appli
C 26	30.8	9.9	2408	US-09-317-179-5	Sequence 5, Appli
C 27	30.8	9.9	3412	US-08-903-325-3	Sequence 3, Appli

28	30.6	9.8	2241	4	US-08-256-799-3	Sequence 3, Appli
29	30.6	9.8	2241	4	US-08-462-437-3	Sequence 3, Appli
C 30	30.4	9.8	786	1	US-08-463-090B-2	Sequence 2, Appli
C 31	30.2	9.7	572	4	US-09-508-542-5	Sequence 5, Appli
C 32	30.2	9.7	1232	4	US-09-392-184-4	Sequence 4, Appli
C 33	30.2	9.7	2473	2	US-08-540-804-3	Sequence 3, Appli
C 34	30.2	9.7	2473	2	US-08-218-265-3	Sequence 3, Appli
C 35	30.2	9.7	2473	3	US-08-521-872-3	Sequence 3, Appli
C 36	30.2	9.7	2473	4	US-08-590-399-3	Sequence 3, Appli
C 37	30.2	9.7	3701	4	US-08-845-258-10	Sequence 10, Appl
C 38	30.2	9.7	3701	4	US-08-990-571-10	Sequence 10, Appl
C 39	30.2	9.7	3701	4	US-08-723-142A-10	Sequence 10, Appl
C 40	30.2	9.7	3701	4	US-09-528-784A-10	Sequence 10, Appl
C 41	30.2	9.7	5303	4	US-08-971-395-4	Sequence 4, Appli
C 42	30	9.6	1048	4	US-09-166-205B-63	Sequence 63, Appl
C 43	30	9.6	3308	3	US-08-714-918-68	Sequence 68, Appl
C 44	30	9.6	3308	4	US-09-265-315-68	Sequence 68, Appl
C 45	30	9.6	3308	4	US-09-265-315-68	Sequence 68, Appl

ALIGNMENTS

RESULT 1
US-09-268-992-7/c
; Sequence 7, Application US/09268992
; Patent No. 6342351
; GENERAL INFORMATION:
; APPLICANT: Chen, H.
; APPLICANT: Freimer, N.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CHROMOSOME-18p RELATED DISORDERS
; FILE REFERENCE: 7853-138
; CURRENT APPLICATION NUMBER: US/09/268,992
; CURRENT FILING DATE: 1999-03-16
; EARLIER APPLICATION NUMBER: 09/236,134
; EARLIER FILING DATE: 1999-01-22
; EARLIER APPLICATION NUMBER: 60/106,056
; EARLIER FILING DATE: 1998-10-28
; EARLIER APPLICATION NUMBER: 60/088,312
; EARLIER FILING DATE: 1998-06-05
; EARLIER APPLICATION NUMBER: 60/078,044
; EARLIER FILING DATE: 1998-03-16
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 72604
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified base
; LOCATION: all n positions
; OTHER INFORMATION: n=a, c, g, or t
US-09-268-992-7

Query Match	11.3%	Score 35;	DB 4;	Length 72604;
Best Local Similarity	52.5%	Pred. No. 0.69;		
Matches	74;	Conservative	1;	Mismatches 66; Indels 0; Gaps 0;
QY	121	AGTCGAATATGTAATACATAATACACCATATATACATATATAGCTGTATGCTGTTCTTA	180	
Db	39189	AGGKAATGKTGAAGATAACCTTTTAAAAAGCTCATTTGTTTGAAGTGTGATGTTAAT	39130	
QY	181	TACTATAGTATGCTGTAACCTCATTCACTGAGAAAAAATGGAAGAAATAGCTATTGCT	240	
Db	39129	TATTAATGCTCAATTAACCTGCAAGCTATTCAATGCTCTATAATTTAGAAATTTACTGA	39070	
QY	241	ATCCATATGGGATACAAAAA	261	
Db	39069	TATACTGTGGCATATGAACA	39049	

RESULT 2

US-09-657-474-7/c
; Sequence 7, Application US/09657474
; Patent No. 6399762
; GENERAL INFORMATION:
; APPLICANT: Chen, H.
; APPLICANT: Freiner, N.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; AND TREATING CHROMOSOME-18p RELATED DISORDERS
; FILE REFERENCE: 7853-138
; CURRENT APPLICATION NUMBER: US/09/657,474
; CURRENT FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: 09/268,992
; PRIOR FILING DATE: 1999-03-16
; PRIOR APPLICATION NUMBER: 09/236,134
; PRIOR FILING DATE: 1999-01-22
; PRIOR APPLICATION NUMBER: 60/106,056
; PRIOR FILING DATE: 1998-10-28
; PRIOR APPLICATION NUMBER: 60/088,312
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/078,044
; PRIOR FILING DATE: 1998-03-16
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 72604
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified base
; LOCATION: all n positions
; OTHER INFORMATION: n=a, c, g, or t
US-09-657-474-7

Query Match 11.3%; Score 35; DB 4; Length 72604;
Best Local Similarity 52.5%; Pred. No. 0.69;
Matches 74; Conservative 1; Mismatches 66; Indels 0; Gaps 0;
QY 121 AGTCGAATAATGAATCTACTACATATACACCATATATACAGCTGTGTGTGTTCTA 180
DB 39189 AGGGAATKTTGAAGATAACCTTTTAAAAAGTCATTGTTTGAGTGTGATTAAT 39130
QY 181 TACTATAGTATGACTTCACTTCATCCATTCAGTGAAGAAAAATCGAAGAAATAGCTATTGCT 240
DB 39129 TATTATGTCATTAACATCAAGCTTCAATCTCTATATATTTAGAAATTTACTGAATTT 39070
QY 241 ATCCATATGGGTACAAAAA 261
DB 39069 TATACGTGGCATATGACAA 39049

RESULT 3
US-09-386-493-3/c
; Sequence 3, Application US/09386493
; Patent No. 6262247
; GENERAL INFORMATION:
; APPLICANT: Kaser, Matthew R.
; APPLICANT: Azimzai, Yalda
; APPLICANT: Yue, Henry
; TITLE OF INVENTION: POLYCYCLIC AROMATIC HYDROCARBON INDUCED MOLECULES
; FILE REFERENCE: PB-0011 US
; CURRENT APPLICATION NUMBER: US/09/386,493
; CURRENT FILING DATE: 1999-08-30
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PERL Program
; SEQ ID NO 3
; LENGTH: 2429
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 2404, 2413, 2423, 2425, 2426
; OTHER INFORMATION: a or g or c or t, unknown, or other
; FEATURE:

; NAME/KEY:
; OTHER INFORMATION: 253053
; PUBLICATION INFORMATION:
US-09-386-493-3

Query Match 10.6%; Score 33; DB 4; Length 2429;
Best Local Similarity 51.7%; Pred. No. 0.62;
Matches 75; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 135 ATACTACAATATACACCATATATACAGCTGTGTGTGTTCTATATAGTATTGATTG 194
DB 2243 AAACATAAATATTAAACGAAAAATTAAGCTGAAGGTGTTTACCTAGTGTCCATAAAGC 2184
QY 195 ACTTGAACTCCATTTCAGTGAAGAAAAATGAAGAATTAAGCTATTGTTATCCATATGGGATA 254
DB 2183 ATGGGTCTCTTTTATTATTAGAAAAATAAAAAACGGCTTTAACACCCCATTTAGGAATA 2124
QY 255 CAAAAAGCAGGGTAACAAAGAAT 279
DB 2123 TAAATAAAATCAACTGAAATATT 2099

RESULT 4
US-09-402-929-3
; Sequence 3, Application US/09402929
; Patent No. 6410825
; GENERAL INFORMATION:
; APPLICANT: Temple University - Of The Commonwealth System of Higher Education
; APPLICANT: Toscani, Antonio
; APPLICANT: Hatton, Kimi
; APPLICANT: Reddy, E. P.
; TITLE OF INVENTION: A-myb NULL MUTANT TRANSGENIC ANIMALS AND
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEIDEL, GONDA, LAVORNA & MONACO, P.C.
; STREET: Suite 1800 Two Penn Center Plaza
; CITY: Philadelphia
; STATE: PA
; COUNTRY: U.S.A.
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/402,929
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US98/06896
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 6056-214 PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5889 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-402-929-3

Query Match 10.6%; Score 33; DB 4; Length 5889;
Best Local Similarity 58.8%; Pred. No. 0.93;
Matches 57; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
QY 203 TCCATTTCAGTGAAAAAATGAAGAATTAAGCTATTGTTATCCATATGGGATACAAAAAG 262

Db 85 TGCATCTTTTAAAAAATCAATGATATGAACAATTTGTTGATTGGGATTGACTGAA 144
QY 263 CAGGTAACAAGAAATCTACATCATCTTGCATTG 299
Db 145 CAGTGAATAGAGGAATTTGTACATACTGACTTTAG 181

RESULT 5

US-08-361-467B-4
; Sequence 4, Application US/08361467B
; Patent No. 5633441
; GENERAL INFORMATION:
; APPLICANT: De Greef, Willy
; APPLICANT: Van Emmelo, John
; APPLICANT: De Oliveria, Dulce E.
; APPLICANT: De Souza, Maria-Helena
; APPLICANT: Van Montagu, Marc
; TITLE OF INVENTION: PLANTS WITH MODIFIED FLOWERS, SEEDS OR
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/361,467B
; FILING DATE: 22-DEC-1994
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/681,492
; FILING DATE: 04-APR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/EP90/01275
; FILING DATE: 01-AUG-1990
; APPLICATION NUMBER: EP 89 402 224.3
; FILING DATE: 04-AUG-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Schulman, Robert M.
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1046 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cdna to mRNA
; IMMEDIATE SOURCE:
; CLONE: 3C9
US-08-361-467B-4

Query Match 10.5%; Score 32.6; DB 1; Length 1046;
Best Local Similarity 51.7%; Pred. No. 0.57;
Matches 74; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 136 TACTACATATACACCATATATAGTACTGTGTCTTCTATCTACTAGTACTGGA 195
Db 860 TATTATAAATGAGCAAGGAAACGACGTAGTTTCTGTCTTCTATCTACTAGTACTGGA 919
QY 196 CTTGAACCTCCATTCAGTGAAGAAATGGAAGAAATAGCTATTGTTGATCCATATGGGATAC 255

Db 920 CTCATATCAAAACCCCAACAGTAAATGATGATTATAAGTTTTCTTCAAAAAAAAAA 979
QY 256 AAAAAGCAGGTTACAAAAGAA 278
Db 980 AAAAAAAAAAAAAAAAAAAAAA 1002

RESULT 6

US-08-484-332C-4
; Sequence 4, Application US/08484332C
; Patent No. 5767374
; GENERAL INFORMATION:
; APPLICANT: De Greef, Willy
; APPLICANT: Van Emmelo, John
; APPLICANT: De Oliveria, Dulce E.
; APPLICANT: De Souza, Maria-Helena
; APPLICANT: Van Montagu, Marc
; TITLE OF INVENTION: PLANTS WITH MODIFIED FLOWERS, SEEDS OR
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,332C
; FILING DATE: 7-JUNE-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/361,467
; FILING DATE: 22-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/681,492
; FILING DATE: 04-APR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/EP90/01275
; FILING DATE: 01-AUG-1990
; APPLICATION NUMBER: EP 89 402 224.3
; FILING DATE: 04-AUG-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Schulman, Robert M.
; REGISTRATION NUMBER: 31,196
; REFERENCE/DOCKET NUMBER: 010830-093
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1046 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cdna to mRNA
; IMMEDIATE SOURCE:
; CLONE: 3C9
US-08-484-332C-4

Query Match 10.5%; Score 32.6; DB 1; Length 1046;
Best Local Similarity 51.7%; Pred. No. 0.57;
Matches 74; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 136 TACTACATATACACCATATATAGTACTGTGTCTTCTATCTACTAGTACTGGA 195

Db 860 TATTAATAAATGAAGCAAGAGAGAAAGCTAGTTTCTTGTGTTCTTCTATTTTGTCTTCT 919
QY 196 CTTGAATCCATTCACTGAGAAAAATGGAAGAAATTAGCTATTGTGTTATCCATATGGGATAC 255
Db 920 CTCTATCAAAACCCACACAGTAAGATGATTTATAGTTTCTTCAAAAAAAGAAAAA 979
QY 256 AAAAAAGCAGGGTAACAAAAAGAA 278
Db 980 AAAAAAAGAAAAAAGAAAAA 1002

RESULT 7
US-08-484-993B-40/c
; Sequence 40, Application US/08484993B
; Patent No. 5837497
; GENERAL INFORMATION:
; APPLICANT: Harris Ph.D., Jeffrey D.
; APPLICANT: Hsu, Kuang T.
; APPLICANT: Podolski, Joseph S.
; TITLE OF INVENTION: Materials and Methods for Immunocontraception
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,993B
; FILING DATE: 09-NOV-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/012,990
; FILING DATE: 29-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/973,341
; FILING DATE: 09-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Clough, David W.
; REGISTRATION NUMBER: 36,107
; REFERENCE/DOCKET NUMBER: 31745
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6653
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1701 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1698
US-08-484-993B-40

Query Match 10.5%; Score 32.6; DB 2; Length 1701;
Best Local Similarity 53.5%; Pred. No. 0.71; Indels 0; Gaps 0;
Matches 68; Conservative 0; Mismatches 59;

QY 3 TCCAGCTCTGTGCAAGAGATGCTGGAATGTCACAGCATCGTATAGCAAAAGCAT 62
Db 840 TACTGCTGGTCTCCAGTATCTGCTTGTGTCACAGGAAGTAATGGAAGTGGAA 781
QY 63 ATTGGCAACAGCTTGATGCCAGCAGAGAGGCCCAATGTGATTCATATTACTAG 122

Db 780 CAGACAAAAAGCTGTGTTGCCCATCACAGGTTTACAGCAGCTGTCTATTCCTAAGGCCAA 721
QY 123 TCGAATA 129
Db 720 GCGCACA 714

RESULT 8
US-08-484-158B-40/c
; Sequence 40, Application US/08484158B
; Patent No. 5976545
; GENERAL INFORMATION:
; APPLICANT: Harris Ph.D., Jeffrey D.
; APPLICANT: Hsu, Kuang T.
; APPLICANT: Podolski, Joseph S.
; TITLE OF INVENTION: Pharmaceutical Compositions for
; NUMBER OF SEQUENCES: 61
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,158B
; FILING DATE: 07-JUNE-95
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/149,223
; FILING DATE: 09-NOV-93
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/012,990
; FILING DATE: 29-JAN-93
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/973,341
; FILING DATE: 09-NOV-92
; ATTORNEY/AGENT INFORMATION:
; NAME: Clough, David W.
; REGISTRATION NUMBER: 36,107
; REFERENCE/DOCKET NUMBER: 32794
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6653
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1701 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1698
US-08-484-158B-40

Query Match 10.5%; Score 32.6; DB 2; Length 1701;
Best Local Similarity 53.5%; Pred. No. 0.71; Indels 0; Gaps 0;
Matches 68; Conservative 0; Mismatches 59;

QY 3 TCCAGCTCTGTGCAAGAGATGCTGGAATGTCACAGCATCGTATAGCAAAAGCAT 62
Db 840 TACTGCTGGTCTCCAGTATCTGCTTGTGTCACAGGAAGTAATGGAAGTGGAA 781
QY 63 ATTGGCAACAGCTTGATGCCAGCAGAGAGGCCCAATGTGATTCATATTACTAG 122

Db 780 CAGAACAAAGCTTGTGTGCATCACAGGTTACACGCGTGTCTTCCTTAAGGCCCAA 721
123 TCGAATA 129
Db 720 GCGCAC 714

RESULT 9
US-08-484-596A-40/c
; Sequence 40, Application US/08484596A
; Patent No. 5981228
; GENERAL INFORMATION:
; APPLICANT: Harris Ph.D., Jeffrey D.
; APPLICANT: Hsu, Kuang T.
; APPLICANT: Podolski, Joseph S.
; TITLE OF INVENTION: Materials and Methods for Immunocontraception
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,596A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/149,223
; FILING DATE: 11-NOV-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/973,341
; FILING DATE: 09-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Clough, David W.
; REGISTRATION NUMBER: 36,107
; REFERENCE/DOCKET NUMBER: 31745
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6653
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1701 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1698
US-08-484-596A-40

Query Match 10.5%; Score 32.6; DB 2; Length 1701;
Best Local Similarity 53.5%; Pred. No. 0.71;
Matches 68; Conservative 0; Mismatches 59; Indels 0; Gaps 0;
QY 3 TCCACTCTGTGTGCAAGAGATGCTGGATGTCACAGCATCGTATAGCAAGCAT 62
Db 840 TACTCTCGGTCTCCAGTGATCTGCTTGTGTGTCACAGGAAGTAAATGGAACCTGGA 781
QY 63 ATTGCAACAGCTTGGATGGCCAGCAGAGAGCCCAATGTGTGATTCATATTCACCTAG 122
Db 780 CAGAACAAAGCTTGTGTGCATCACAGGTTACACGCGTGTCTTCCTTAAGGCCCAA 721
QY 123 TCGAATA 129
Db 720 GCGCAC 714

Db 720 GCGCAC 714
RESULT 10
US-08-480-150A-40/c
; Sequence 40, Application US/08480150A
; Patent No. 5989550
; GENERAL INFORMATION:
; APPLICANT: Harris Ph.D., Jeffrey D.
; APPLICANT: Hsu, Kuang T.
; APPLICANT: Podolski, Joseph S.
; TITLE OF INVENTION: Materials and Methods for Immunocontraception
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,150A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/149,223
; FILING DATE: 09-NOV-1993
; APPLICATION NUMBER: 08/012,990
; FILING DATE: 29-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/973,341
; FILING DATE: 09-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Clough, David W.
; REGISTRATION NUMBER: 36,107
; REFERENCE/DOCKET NUMBER: 31745
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6653
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1701 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1698
US-08-480-150A-40

Query Match 10.5%; Score 32.6; DB 2; Length 1701;
Best Local Similarity 53.5%; Pred. No. 0.71;
Matches 68; Conservative 0; Mismatches 59; Indels 0; Gaps 0;
QY 3 TCCACTCTGTGTGCAAGAGATGCTGGATGTCACAGCATCGTATAGCAAGCAT 62
Db 840 TACTCTCGGTCTCCAGTGATCTGCTTGTGTGTCACAGGAAGTAAATGGAACCTGGA 781
QY 63 ATTGCAACAGCTTGGATGGCCAGCAGAGAGCCCAATGTGTGATTCATATTCACCTAG 122
Db 780 CAGAACAAAGCTTGTGTGCATCACAGGTTACACGCGTGTCTTCCTTAAGGCCCAA 721
QY 123 TCGAATA 129
Db 720 GCGCAC 714

RESULT 11

US-08-458-731-40/C
; Sequence 40, Application US/08458731
; Patent No. 6001599
; GENERAL INFORMATION:
; APPLICANT: Harris Ph.D., Jeffrey D.
; APPLICANT: Hsu, Kuang T.
; APPLICANT: Podolski, Joseph S.
; TITLE OF INVENTION: Materials and Methods for Immunocontraception
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/458,731
; FILING DATE: 09-NOV-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/012,990
; FILING DATE: 29-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/973,341
; FILING DATE: 09-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Clough, David W.
; REGISTRATION NUMBER: 36,107
; REFERENCE/DOCKET NUMBER: 31745
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6653
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1701 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; NAME/KEY: CDS
; LOCATION: 1..1698
; US-08-458-731-40

Query Match 10.5%; Score 32.6; DB 3; Length 1701;
Best Local Similarity 53.5%; Pred. No. 0.71;
Matches 68; Conservative 0; Mismatches 59; Indels 0; Gaps 0;
QY 3 TCCAGCTCTGTGCAAGGAGATGCTGGAATGTCACAGCATCGTATAGCAAGAGCAT 62
Db 840 TACTGCTCGGTCTCCAGTATCTGTTGGTGCCACAGGAGTAATGGAACTGGAA 781
QY 63 ATTGGCAACAGCTTGGATGCCAGCAGAGAGGCCCAAAATGTGTGATTTCATTTACTAG 122
Db 780 CAGAACAAAGCTGTGTTGCCATCACAGGTTACAGGCACGTCTCATTCTTAAGGCCAA 721
QY 123 TCGAATA 129
Db 720 GCGCACA 714

RESULT 12

US-08-149-223A-40/c

; Sequence 40, Application US/08149223A
; Patent No. 6027727
; GENERAL INFORMATION:
; APPLICANT: Harris Ph.D., Jeffrey D.
; APPLICANT: Hsu, Kuang T.
; APPLICANT: Podolski, Joseph S.
; TITLE OF INVENTION: Materials and Methods for Immunocontraception
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/149,223A
; FILING DATE: 09-NOV-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/012,990
; FILING DATE: 29-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/973,341
; FILING DATE: 09-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Clough, David W.
; REGISTRATION NUMBER: 36,107
; REFERENCE/DOCKET NUMBER: 31745
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6653
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1701 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; NAME/KEY: CDS
; LOCATION: 1..1698
; US-08-149-223A-40

Query Match 10.5%; Score 32.6; DB 3; Length 1701;
Best Local Similarity 53.5%; Pred. No. 0.71;
Matches 68; Conservative 0; Mismatches 59; Indels 0; Gaps 0;
QY 3 TCCAGCTCTGTGCAAGGAGATGCTGGAATGTCACAGCATCGTATAGCAAGAGCAT 62
Db 840 TACTGCTCGGTCTCCAGTATCTGTTGGTGCCACAGGAGTAATGGAACTGGAA 781
QY 63 ATTGGCAACAGCTTGGATGCCAGCAGAGAGGCCCAAAATGTGTGATTTCATTTACTAG 122
Db 780 CAGAACAAAGCTGTGTTGCCATCACAGGTTACAGGCACGTCTCATTCTTAAGGCCAA 721
QY 123 TCGAATA 129
Db 720 GCGCACA 714

RESULT 13

US-09-453-702B-64/c
; Sequence 64, Application US/09453702B
; Patent No. 6365723
; GENERAL INFORMATION:
; APPLICANT: Blattner, Frederick R.

```

; Burland, Valerie
; Perna, Nicole T.
; Plunkett, Guy
; Welch, Rod
;
; TITLE OF INVENTION: No. 6365723e1 Sequences of E. coli O157
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch. 1.44mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 8.0
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/453,702B
; FILING DATE: 03-Dec-1999
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/110,955
; FILING DATE: 04-DEC-1998
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 960296.95017
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 251-5000
; TELEFAX: (608) 251-9166
;
; INFORMATION FOR SEQ ID NO: 64:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31728
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 64:
US-09-453-702B-64

Query Match 10.5%; Score 32.6; DB 4; Length 31728;
Best Local Similarity 50.3%; Pred. No. 2.7;
Matches 80; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 110 TCATATCACTAGTCGAATAATGTAATCACTACATATACACCATATATAGTGTAT 169
DB 3236 TTAATCTTAAATAACAGATAATTCAGATTGCAATTTAGGAAATATATAGTGAC 3177
QY 170 GTGTTGCTCTACTATAGTGTGACTTGAACTCCATTGCAAGTGAATAAATGGAAGAT 229
DB 3176 TTAATGTTCTGCGCTTTATTTCTGAGTAATCTGATTTGTAATACACACTAAGGTAT 3117
QY 230 TAGCTATTTGATCCATATGGGATACAAAAAGCAGGT 268
DB 3116 AAAATGTAATGATCTACATATATAAATGATGAAGTAAGAT 3078

RESULT 14
US-09-426-290-1/c
; Sequence 1, Application US/09426290
; Patent No. 6410712
; GENERAL INFORMATION:
; APPLICANT: Berglind Ran Olafsdottir
; APPLICANT: Jeffrey Gulcher
; TITLE OF INVENTION: HUMAN NARCOLEPSY GENE
; FILE REFERENCE: 2345.2001-000
; CURRENT APPLICATION NUMBER: US/09/426,290
; CURRENT FILING DATE: 1999-10-25
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1

```

```

; LENGTH: 168575
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (21181)...(21403)
; NAME/KEY: CDS
; LOCATION: (95252)...(95430)
; NAME/KEY: CDS
; LOCATION: (101753)...(101996)
; NAME/KEY: CDS
; LOCATION: (110324)...(110439)
; NAME/KEY: CDS
; LOCATION: (124058)...(124278)
; NAME/KEY: CDS
; LOCATION: (127009)...(127130)
; NAME/KEY: CDS
; LOCATION: (128910)...(129139)
;
; US-09-426-290-1
Query Match 10.5%; Score 32.6; DB 4; Length 168575;
Best Local Similarity 51.0%; Pred. No. 5.7;
Matches 77; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 150 CCATATATACAGACTGTATGTGTGTTCTTCTATATCTATATAGTGTGACTTGAACCTCATTC 209
DB 14689 CAATCACTACTTGTGTATCAGGCGCTGAACCTCAATTTGAAGGGAGTAATTCCTTCTTCAAAAC 14630
QY 210 AGTGAATAAATGGAGAATTAGCTATTTGTATTCATATGATGGATACAAAAAGCAGGTA 269
DB 14629 AGTGAATATTTTGTGAAAAATATACATATGTGCCGCCCTTTGTTATGTAACACCAAGGAA 14570
QY 270 ACAAAAGAATCTACATCATCTTGCCATTGC 300
DB 14569 ATTCCTGGATATGATATTTCTTGCCTATTAC 14539

RESULT 15
US-08-765-381-3
; Sequence 3, Application US/08765381
; Patent No. 6083724
; GENERAL INFORMATION:
; APPLICANT: Commonwealth Scientific and Industrial Research Organisation
; TITLE OF INVENTION: No. 6083724e1 avian cytokines and genetic
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully Scott Murphy and Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City, New York
; STATE: New York
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/765,381
; FILING DATE: 19-DEC-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PN1542/95
; FILING DATE: 06-MAR-1995
; APPLICATION NUMBER: PCT/AU96/00114
; FILING DATE: 05-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: PRESSER, LEOPOLD
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 1-516-742-4343
; TELEFAX: 1-516-742-4366
; INFORMATION FOR SEQ ID NO: 3:

```

